



Db	241		GGGTGGCGCTCGCTCCAGCGCCGGCGCGGGAGCGCGGGCGCGGTGGCGCGCC	300
Qy	301	CCGGACCGTATCCCTCCGGCGCCCTCCCGCGCCCGCCCGCGCCCTCCCTCCCGG	360	
Db	301	CCGGACCGTATCCCTCCGGCGCCCTCCCGCGCCCGCCCGCGCCCTCCCTCCCGG	360	
Qy	361	CAGAGCTCCCTCCCTCCGGCTCAGACTGTGTTTGGTAGCAACGCGCAACGCGGGCGCG	420	
Db	361	CAGAGCTCCCTCCCTCCGGCTCAGACTGTGTTTGGTAGCAACGCGCAACGCGGGCGCG	420	
Qy	421	TTTTCGGCCCGCTCCGGCGGCTCTTGCTCTCGCGGCGCTCCCGGCCCTTCGTGTC	480	
Db	421	TTTTCGGCCCGCTCCGGCGGCTCTTGCTCTCGCGGCGCTCCCGGCCCTTCGTGTC	480	
Qy	481	GTCTTCTCCCTCCCGCAGCCCGGCGCCCTCCGGCGCGCCCAACCCGGCGCTCCCG	540	
Db	481	GTCTTCTCCCTCCCGCAGCCCGGCGCCCTCCGGCGCGCCCAACCCGGCGCTCCCG	540	
Qy	541	CTCGCGCCCTGCGTCCCGCCCGGCTTCGCGGTCTCTTTGGGCGCCCGGCTCCCGC	600	
Db	541	CTCGCGCCCTGCGTCCCGCCCGGCTTCGCGGTCTCTTTGGGCGCCCGGCTCCCGC	600	
Qy	601	TGTCCCGCCCGGTGCGAGCGGTGTATGGGCCCTTCAACATGTCTGCTGAAGCCCG	660	
Db	601	TGTCCCGCCCGGTGCGAGCGGTGTATGGGCCCTTCAACATGTCTGCTGAAGCCCG	660	
Qy	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720	
Db	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720	
Qy	721	CAGCGCGCCCGGCGTGCCTTCGCAAGCCCGGCGCGCGCGCTTCTAGCGTCG	780	
Db	721	CAGCGCGCCCGGCGTGCCTTCGCAAGCCCGGCGCGCGCTTCTAGCGTCG	780	
Qy	781	CCCGCGCGCGCTTCGCGTCTCGTCTCGGTCTCTCGTCTCGGCGCGCGCTCC	840	
Db	781	CCCGCGCGCGCTTCGCGTCTCGTCTCGGTCTCTCGTCTCGGCGCGCGCTCC	840	
Qy	841	TCCTCGGTGTCGGCGACCTCCGGCGCGGAGCGCCCGGCTGGGCGAGGTCGAAC	900	
Db	841	TCCTCGGTGTCGGCGACCTCCGGCGCGGAGCGCCCGGCTGGGCGAGGTCGAAC	900	
Qy	901	AGTAACAAAGGACTGCCTCAGCTACAGATTTCTTTTGATGGAATCATGCAAAATAGG	960	
Db	901	AGTAACAAAGGACTGCCTCAGCTACAGATTTCTTTTGATGGAATCATGCAAAATAGG	960	
Qy	961	ATGGTTCAATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTCAAGTGAAATATGA	1020	
Db	961	ATGGTTCAATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTCAAGTGAAATATGA	1020	
Qy	1021	GGTATATGAAGGATTTTAAACCTTACAGTCCGAGTGTGATTTGGTACTTGATGCC	1080	
Db	1021	GGTATATGAAGGATTTTAAACCTTACAGTCCGAGTGTGATTTGGTACTTGATGCC	1080	
Qy	1081	GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGGAAACCTGAAGAAATATGAGAGT	1140	
Db	1081	GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGGAAACCTGAAGAAATATGAGAGT	1140	
Qy	1141	ATTTTGTTCAAATGTTTACAGCTTCTGCTATCAGTCTGCTTAAAGATATGCAATCCAGTTAT	1200	
Db	1141	ATTTTGTTCAAATGTTTACAGCTTCTGCTATCAGTCTGCTTAAAGATATGCAATCCAGTTAT	1200	
Qy	1201	GCAAAAGAGATGCTTTTACCTGCTATCAGTCTGCTTAAAGATATGCAATCCAGTTAT	1260	
Db	1201	GCAAAAGAGATGCTTTTACCTGCTATCAGTCTGCTTAAAGATATGCAATCCAGTTAT	1260	
Qy	1261	GAGAAGGACCTGGAGCCCTGGGATGCAAGTCACTCAGAGCAATGAGAACTTGAGCT	1320	
Db	1261	GAGAAGGACCTGGAGCCCTGGGATGCAAGTCACTCAGAGCAATGAGAACTTGAGCT	1320	
Qy	1321	TTGAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGAA	1380	

Db	1321	TTGGAAAATGACGTATCTAAATGGATGGGATCCCAATGATATGTTCTCGATATAATAGAGAA	1380
Qy	1381	AATTTATCGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA	1440
Db	1381		
Db	1381	AATTTATCGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA	1440
Qy	1441	AGAGATAACTCAGAGAAGATTTTTAAAAACGGGAAGCAAGGGCAAAACCAAGTTAGCAGAGA	1500
Db	1441	AGAGATAACTCAGAGAAGATTTTTAAAAACGGGAAGCAAGGGCAAAACCAAGTTAGCAGAGA	1500
Qy	1501	ATTGAGTCAGTGCAGTGAACAAGCTCGAGTGGCCCTGGAAAAATGATGATAGGAGTGAG	1560
Db	1501	ATTGAGTCAAGTGCAGTGAACAAGCTCGAGTGGCCCTGGAAAAATGATGATAGGAGTGAG	1560
Qy	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACCGTGAAGGGCACACATAAAC	1620
Db	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACCGTGAAGGGCACACATAAAC	1620
Qy	1621	ACTAGGGAAAAATAATATATTTCTCTCTGGACAAGAAATAGAGAAGTCATATCCTGGGGA	1680
Db	1621	ACTAGGGAAAAATAATATATTTCTCTCTGGACAAGAAATAGAGAAGTCATATCCTGGGGA	1680
Qy	1681	AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
Db	1681	AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
Qy	1741	TCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGGTTCAGACCAAGAGTAGTTAAT	1800
Db	1741	TCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGGTTCAGACCAAGAGTAGTTAAT	1800
Qy	1801	GGAGGTGTTCCCTGGCCCATCGCCTTGCCCATCTCCTCTCGCCGCCACCTCTCGCTAC	1860
Db	1801	GGAGGTGTTCCCTGGCCCATCGCCTTGCCCATCTCCTCTCGCCGCCACCTCTCGCTAC	1860
Qy	1861	CAGTCAGGTCCCAACTCTTTCCACCTCGGGCAGCCACCCCTACAGGGCCGCTCCAGG	1920
Db	1861	CAGTCAGGTCCCAACTCTTTCCACCTCGGGCAGCCACCCCTACAGGGCCGCTCCAGG	1920
Qy	1921	CCCCCTCGGGGCCATCCAGACCCCGTCTCACCCCTCGCTCATGGTTCTCCAGTCTCT	1980
Db	1921	CCCCCTCGGGGCCATCCAGACCCCGTCTCACCCCTCGCTCATGGTTCTCCAGTCTCT	1980
Qy	1981	GTCTCTACTATGCTAAACGCATGTCTTCAGAAAGGGCTTCCAAAGGATGTCCCCAAAGGCC	2040
Db	1981	GTCTCTACTATGCTAAACGCATGTCTTCAGAAAGGGCTTCCAAAGGATGTCCCCAAAGGCC	2040
Qy	2041	CAGCGACATCTCGAAATCAAGATTTTCGTGGGAGGGTTCCATATCCAGTGGCCTA	2100
Db	2041	CAGCGACATCTCGAAATCAAGATTTTCGTGGGAGGGTTCCATATCCAGTGGCCTA	2100
Qy	2101	GAATTTGTAATCCACACCCACCCAGTGAAGCAGCTACTCCTCCAGTGAAGGACCAAGT	2160
Db	2101	GAATTTGTAATCCACACCCACCCAGTGAAGCAGCTACTCCTCCAGTGAAGGACCAAGT	2160
Qy	2161	CCCTCGGGGGAAACGTGGTCAATCAGTGGTCAAGTGGGTTCCAAAGATTATCCCCCTAAA	2220
Db	2161	CCCTCGGGGGAAACGTGGTCAATCAGTGGTCAAGTGGGTTCCAAAGATTATCCCCCTAAA	2220
Qy	2221	CATAGACCCAGGTCTCCACAGACAAACAGTATTTGGAAATACCCCCCAGTGGGCCAGTTC	2280
Db	2221	CATAGACCCAGGTCTCCACAGACAAACAGTATTTGGAAATACCCCCCAGTGGGCCAGTTC	2280
Qy	2281	GCTTCTCCCCACAGCTGGTATTTATTCACACTGAAGCTGTTGGCAATGCCATTTCCAGCTG	2340
Db	2281	GCTTCTCCCCACAGCTGGTATTTATTCACACTGAAGCTGTTGGCAATGCCATTTCCAGCTG	2340
Qy	2341	TCTCTACGCTGTAGTCTGTGATCGAACAGAGCTGTTACCCCTTCTAGTGAGGCTAAA	2400
Db	2341	TCTCTACGCTGTAGTCTGTGATCGAACAGAGCTGTTACCCCTTCTAGTGAGGCTAAA	2400
Qy	2401	GATTCAGGCTTCAAGATCAGAGCGCAACTCTCTCGAGGGAATAAGAAAAATATTTAAA	2460
Db	2401	GATTCAGGCTTCAAGATCAGAGCGCAACTCTCTCGAGGGAATAAGAAAAATATTTAAA	2460



ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 4481)
TITLE	Brahmachari,S.K., Choudhry,S., Mukerji,M. and Jain,S.
JOURNAL	Method of detection of allelic variants of SCA2 gene
	Patent: US 623927-A 20 23-SEP-2003;
	Council of Scientific and Industrial Research; New Delhi;
INX;	
FEATURES	Location/Qualifiers
source	1..4481
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Query Match	100.0%; Score 4481; DB 6; Length 4481;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 4481; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ACCCCGAGAAAGCAACCCAGGGCGCCCGCTCCTCAGTGTGCTTCCCGGCCCGGG 60
Db	1 ACCCCGAGAAAGCAACCCAGGGCGCCCGCTCCTCAAGTGTGCTTCCCGGCCCGGG 60
Qy	61 GCCACCTCAGCTTCGTCTTCGGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
Db	61 GCCACCTCAGCTTCGTCTTCGGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
Qy	121 CACCTCCGCTCCCAACCCCGCGCTCGGCGCGCCCGCTCCGATGCGCTCAGCGCGCGCA 180
Db	121 CACCTCCGCTCCCAACCCCGCGCTCGGCGCGCCCGCTCCGATGCGCTCAGCGCGCGCA 180
Qy	181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTGCGCGCTTCGCGCGACGAGTGGGCC 240
Db	181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTGCGCGCTTCGCGCGACGAGTGGGCC 240
Qy	241 GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
Db	241 GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
Qy	301 CCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db	301 CCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy	361 CAGAGTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGGCGGCGG 420
Db	361 CAGAGTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGGCGGCGG 420
Qy	421 TTTGGCGCGGCTCCGCGCGCTCTGCTGCTCGGCGGCGCTCCCGCGCGCTTCGTGTC 480
Db	421 TTTGGCGCGGCTCCGCGCGCTCTGCTGCTCGGCGGCGCTCCCGCGCGCTTCGTGTC 480
Qy	481 GTCTTTCTCCCTCCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db	481 GTCTTTCTCCCTCCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Qy	541 CTCGGCGCGGCTCCCGCGCGGTTCCGCGCTCTTGGCGCGCGCGCGCGCGCGCG 600
Db	541 CTCGGCGCGGCTCCCGCGCGGTTCCGCGCTCTTGGCGCGCGCGCGCGCGCGCG 600
Qy	601 TGTCCCGCGCGGCTCGAGCGGCTGATGGCGCGCTCACCATGCTGAGCGCGCGCG 660
Db	601 TGTCCCGCGCGGCTCGAGCGGCTGATGGCGCGCTTACCATGCTGAGCGCGCGCG 660
Qy	661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Db	661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Qy	721 CAGCGCGCGCGCGGCTGCGCAATGTCCGAAGCCCGCGCGCGCGCGCGCGCGCT 780
Db	721 CAGCGCGCGCGCGGCTGCGCAATGTCCGAAGCCCGCGCGCGCGCGCGCGCGCT 780
Qy	781 CCGCGCGCGCGGCTTCGCGGCTCGCTCGGCTCTCTGCTCTCGGCGCGCGCGCTCC 840
Db	781 CCGCGCGCGCGGCTTCGCGGCTCTCTGCTCTCTGCTCTCGGCGCGCGCGCTCC 840

Qy	841 TCCTCGGTGTCGCGCGACCTCCGCGCGCGGAGGCCCGGCTTGGCAGAGTTCGAAAC 900
Db	841 TCCTCGGTGTCGCGCGACCTCCGCGCGCGGAGGCCCGGCTTGGCAGAGTTCGAAAC 900
Qy	901 AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCANATATGAGG 960
Db	901 AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCANATATGAGG 960
Qy	961 ATGGTTCATATACCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTCAAAATATGGA 1020
Db	961 ATGGTTCATATACCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTCAAAATATGGA 1020
Qy	1021 GGTATATGAAGGAGTTTTTAAACCTTACAGTCCGAAAGTGTGATTTGGTACTTTGATGCC 1080
Db	1021 GGTATATGAAGGAGTTTTTAAACCTTACAGTCCGAAAGTGTGATTTGGTACTTTGATGCC 1080
Qy	1081 GCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGTGAAGAAATAATGGAGAGT 1140
Db	1081 GCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGTGAAGAAATAATGGAGAGT 1140
Qy	1141 ATTTTGTTCAAATGTTTTCAGACTTTTGTGTGATCAGTTTAAAGATATGGACTCCAGTTAT 1200
Db	1141 ATTTTGTTCAAATGTTTTCAGACTTTTGTGTGATCAGTTTAAAGATATGGACTCCAGTTAT 1200
Qy	1201 GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCAACACAAA 1260
Db	1201 GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCAACACAAA 1260
Qy	1261 GAGAAGACCTGGAGCCCTGGGATGAGGTGAACTACACAGCAATGAGGACTTTGAGGCT 1320
Db	1261 GAGAAGACCTGGAGCCCTGGGATGAGGTGAACTACACAGCAATGAGGACTTTGAGGCT 1320
Qy	1321 TTGAAATGACGTATCTAATGATGGATGCCAATGATATGTTTCGATATAATGAAGAA 1380
Db	1321 TTGAAATGACGTATCTAATGATGGATGCCAATGATATGTTTCGATATAATGAAGAA 1380
Qy	1381 AATTATCGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTGA 1440
Db	1381 AATTATCGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTGA 1440
Qy	1441 AGAGATACTCAGAAAGATTTTAAACCGGAGCAAGGCAACACAGTTAGCAGAGAA 1500
Db	1441 AGAGATACTCAGAAAGATTTTAAACCGGAGCAAGGCAACACAGTTAGCAGAGAA 1500
Qy	1501 ATTGAGTCAAGTGCCTCAGTCAAAAGCTCGAGTGGCCCTGGAAAAATGATAGAGGTGAG 1560
Db	1501 ATTGAGTCAAGTGCCTCAGTCAAAAGCTCGAGTGGCCCTGGAAAAATGATAGAGGTGAG 1560
Qy	1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGAGTGAAGTGAAGGGGACACAGATAAAC 1620
Db	1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGAGTGAAGTGAAGGGGACACAGATAAAC 1620
Qy	1621 ACTAGGAAAAATTAATATATTTCTCTGGACAAAGAAATAGAGAGTCAATATCTCGGGA 1680
Db	1621 ACTAGGAAAAATTAATATATTTCTCTGGACAAAGAAATAGAGAGTCAATATCTCGGGA 1680
Qy	1681 AGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCTCGGCTCCATGCGGCTCAAGA 1740
Db	1681 AGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCTCGGCTCCATGCGGCTCAAGA 1740
Qy	1741 TCCACTTCTCACCTTCAGATTTCAACCCGAAATTCAGATTCAGACCAAGAGTAGTTAAT 1800
Db	1741 TCCACTTCTCACCTTCAGATTTCAACCCGAAATTCAGATTCAGACCAAGAGTAGTTAAT 1800
Qy	1801 GGAGGTGTTCCTGGCCATCGCTTGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Db	1801 GGAGGTGTTCCTGGCCATCGCTTGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Qy	1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGCGAGCCACCCCTACACGGCCGCTCCAGG 1920
Db	1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGCGAGCCACCCCTACACGGCCGCTCCAGG 1920



QY 1921 CCCCCCTCGGGCATCCAGACCCCGCTCTCAGCCCTCTGCTCATGGTTCTCCAGCTCT 1980  
DB 1921 CCCCCCTCGGGCATCCAGACCCCGCTCTCAGCCCTCTGCTCATGGTTCTCCAGCTCT 1980  
QY 1981 GTCTCTACTATGCTTAAACGCAATGCTCTCAGAGGGGCTTCCAGAGGATGTCCTCCAAAGGCC 2040  
DB 1981 GTCTCTACTATGCTTAAACGCAATGCTCTCAGAGGGGCTTCCAGAGGATGTCCTCCAAAGGCC 2040  
QY 2041 CAGCGACATCTCTCGAAATCACAGAGTTTCTGTGGGAGGGTTCATATCCAGTGGCCTA 2100  
DB 2041 CAGCGACATCTCTCGAAATCACAGAGTTTCTGTGGGAGGGTTCATATCCAGTGGCCTA 2100  
QY 2101 GAATTTGTATCCACAGACCCACCCAGTGAAGCAGTCTCTCCAGTAGCAGAGCCAGT 2160  
DB 2101 GAATTTGTATCCACAGACCCACCCAGTGAAGCAGTCTCTCCAGTAGCAGAGCCAGT 2160  
QY 2161 CCCTCGGGGGAAAGTGGTCTCAGTGGTTCAGTGGGGTTCAGAGTATTCCTCTAAACT 2220  
DB 2161 CCCTCGGGGGAAAGTGGTCTCAGTGGTTCAGTGGGGTTCAGAGTATTCCTCTAAACT 2220  
QY 2221 CATPAGACCCAGGTCTCCAGACAGAAACAGTATTTGGAAATACCCCCAGTGGGCGAGTTCT 2280  
DB 2221 CATPAGACCCAGGTCTCCAGACAGAAACAGTATTTGGAAATACCCCCAGTGGGCGAGTTCT 2280  
QY 2281 GCTTCTCCAGAGCTGGTATTTTCCAACTGAAGCTGTGGCAGTGCCTATTCAGCTGCA 2340  
DB 2281 GCTTCTCCAGAGCTGGTATTTTCCAACTGAAGCTGTGGCAGTGCCTATTCAGCTGCA 2340  
QY 2341 TCTCTACGCTCTAGTCTGATCGATCGACAGAGCTGTACCCCTCTAGTGGGCTAA 2400  
DB 2341 TCTCTACGCTCTAGTCTGATCGATCGACAGAGCTGTACCCCTCTAGTGGGCTAA 2400  
QY 2401 GATTCAGGGTTCAGATCAGAGGAGAACTCTCTCGCGGGAAATAAGAAATATTTAA 2460  
DB 2401 GATTCAGGGTTCAGATCAGAGGAGAACTCTCTCGCGGGAAATAAGAAATATTTAA 2460  
QY 2461 CCCAATGAACATCACTAGTCTCTCAAAGCTGAAACAAAGGTATATCACAGTTGTT 2520  
DB 2461 CCCAATGAACATCACTAGTCTCTCAAAGCTGAAACAAAGGTATATCACAGTTGTT 2520  
QY 2521 TCTGAACATAGAAACAGATGATGATTTAAAGAAATTTAAGAAATGATTTAGTTACAG 2580  
DB 2521 TCTGAACATAGAAACAGATGATGATTTAAAGAAATTTAAGAAATGATTTAGTTACAG 2580  
QY 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAATAGAGGGAGAAAA 2640  
DB 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAATAGAGGGAGAAAA 2640  
QY 2641 TCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTCTAAGGATTTCTTTTCAATTGAAAT 2700  
DB 2641 TCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTCTAAGGATTTCTTTTCAATTGAAAT 2700  
QY 2701 AGCAGAGCAACTGTACCGAGTGCAGCAGCAGCGAATAGCCCGCATTTCCCTTCA 2760  
DB 2701 AGCAGAGCAACTGTACCGAGTGCAGCAGCAGCGAATAGCCCGCATTTCCCTTCA 2760  
QY 2761 ATACTTAGTAACCGGAGCAACAGAGGGGACCTGAGTCACTTCCCAAGGGGTTACAGCT 2820  
DB 2761 ATACTTAGTAACCGGAGCAACAGAGGGGACCTGAGTCACTTCCCAAGGGGTTACAGCT 2820  
QY 2821 TCCAGGCCAGCATGTAAACAGAGAAAGACGATAAGGAAGAGAAAGACGCGCTGAG 2880  
DB 2821 TCCAGGCCAGCATGTAAACAGAGAAAGACGATAAGGAAGAGAAAGACGCGCTGAG 2880  
QY 2881 CAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTCTCTCTCT 2940  
DB 2881 CAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTCTCTCTCT 2940  
QY 2941 CAGCAAAGCTTCTACTACCCCAACTTCACTCTGGGCTCAAGCACAACCTTAGCCCATCT 3000  
DB 2941 CAGCAAAGCTTCTACTACCCCAACTTCACTCTGGGCTCAAGCACAACCTTAGCCCATCT 3000  
QY 3001 ATGGTGGGTCTATCAACAGGCAACTCCAGTTTATCTCAGCGCTGTGTTTTTGGCACCAAT 3060

DB 3001 ATGGTGGGTCTATCAACAGGCAACTCCAGTTTATCTCAGCGCTGTGTTTTGGCACCAAT 3060  
QY 3061 ATGATGTATCCAGTCCCGAGTGAAGCCGCGTGAACCTTTATACCCAAATACCTATGAGC 3120  
DB 3061 ATGATGTATCCAGTCCCGAGTGAAGCCGCGTGAACCTTTATACCCAAATACCTATGAGC 3120  
QY 3121 CCCATGCCAGTGAATCAAGCCAGACATATAGAGCAGTACCAAAATATGCCCAACAGCGG 3180  
DB 3121 CCCATGCCAGTGAATCAAGCCAGACATATAGAGCAGTACCAAAATATGCCCAACAGCGG 3180  
QY 3181 CAAAGACAGCATCATCAGAGTGCATGATGCAACCCAGCGTTCAGCAGCGGCGCCACCGATT 3240  
DB 3181 CAAAGACAGCATCATCAGAGTGCATGATGCAACCCAGCGTTCAGCAGCGGCGCCACCGATT 3240  
QY 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGGCTACAGTCTCAGCAGTTCCTCA 3300  
DB 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGGCTACAGTCTCAGCAGTTCCTCA 3300  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCTCATCTATAGT 3360  
DB 3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCTCATCTATAGT 3360  
QY 3361 CCTGTAATACAGGGTAATGCTAGAAATGAGGCAACCAACACAGCCCGCAGCTTGTTTA 3420  
DB 3361 CCTGTAATACAGGGTAATGCTAGAAATGAGGCAACCAACACAGCCCGCAGCTTGTTTA 3420  
QY 3421 GTATCTCTTTCAGCAACTCAGTACGGGGCTATGAGCAGACGATGCGATGTATGATGT 3480  
DB 3421 GTATCTCTTTCAGCAACTCAGTACGGGGCTATGAGCAGACGATGCGATGTATGATGT 3480  
QY 3481 CCCAAATTTACATCAACAAAGGAGCAAGCCCTTCTTTTCTACTTTGCGCATTTCCACGGC 3540  
DB 3481 CCCAAATTTACATCAACAAAGGAGCAAGCCCTTCTTTTCTACTTTTGCCATTTCCACGGC 3540  
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTCTCACCCACATACTTCCACACCT 3600  
DB 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTCTCACCCACATACTTCCACACCT 3600  
QY 3601 CAGCCTTCAGCTACCCCCACTGGACAGCAGCAAGAGCAACATGGTGAAGTCACTCTGCA 3660  
DB 3601 CAGCCTTCAGCTACCCCCACTGGACAGCAGCAAGAGCAACATGGTGAAGTCACTCTGCA 3660  
QY 3661 CCCAGTCTCTGTTTCCAGCAGCAGCAGGCGGCGCCAGGCTCTCCATCTGGCAGTCCA 3720  
DB 3661 CCCAGTCTCTGTTTCCAGCAGCAGCAGCAGGCGGCGCCAGGCTCTCCATCTGGCAGTCCA 3720  
QY 3721 CAGCAGCTCAGCCCAATTTTACCAAGCGGGCTTTGGCCAACTCCACCTCCATGACACCT 3780  
DB 3721 CAGCAGCTCAGCCCAATTTTACCAAGCGGGCTTTGGCCAACTCCACCTCCATGACACCT 3780  
QY 3781 GCTCCCAACAGCAGCTCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
DB 3781 GCTCCCAACAGCAGCTCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
QY 3841 ATCCATCTCTTCAAGTTTCCAGCGGGTATACCAACCCACCCACATGGCCAGCTACCT 3900  
DB 3841 ATCCATCTCTTCAAGTTTCCAGCGGGTATACCAACCCACCCACATGGCCAGCTACCT 3900  
QY 3901 CAGGCTCATGTACAGTTCAGGAATGTTTCTCTCATCTCAACTGCCCCATGGCCCAATGATG 3960  
DB 3901 CAGGCTCATGTACAGTTCAGGAATGTTTCTCTCATCTCAACTGCCCCATGGCCCAATGATG 3960  
QY 3961 CTAAATGACACAGCAGCAGCCCGCGGCTCCCGAGCGGCTCTCGCTCAAAAGTGCATACAG 4020  
DB 3961 CTAAATGACACAGCAGCAGCCCGCGGCTCCCGAGCGGCTCTCGCTCAAAAGTGCATACAG 4020  
QY 4021 CCCATTTCCAGTCTCGAACAACAGCGCATTTTCCCTCTATATGACGCAACCTTTTCAAGGCC 4080  
DB 4021 CCCATTTCCAGTCTCGAACAACAGCGCATTTTCCCTCTATATGACGCAACCTTTTCAAGGCC 4080  
QY 4081 CACCACCAACAGCAGTGTAAAGCTCCCTGGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140

Db	4081	CACCACCAACAGCAGTGTGAAGCTGCCCTGGAGGAA	CGGAAAGCCAAATTCCTCCTCCTC	4140
QY	4141	CCTTCTACTGCTTCTACCAACTGGAAGACAGAAACTAGAAATTCATTTATTTGTTTT		4200
Db	4141	CCTTCTACTGCTTCTACCAACTGGAAGACAGAAACTAGAAATTCATTTATTTGTTTT		4200
QY	4201	TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAACAGTTCACATTGCAG		4260
Db	4201	TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAACAGTTCACATTGCAG		4260
QY	4261	TGGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTTCCATAATTTCCA		4320
Db	4261	TGGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTTCCATAATTTCCA		4320
QY	4321	TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGC	CGGAAACTGGAAGTTATTT	4380
Db	4321	TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGC	CGGAAACTGGAAGTTATTT	4380
QY	4381	ATTTTAAATACCTTGAAAGTCATGAACATCAGCTAGCAAAAGAAAGTAACAGAGT		4440
Db	4381	ATTTTAAATACCTTGAAAGTCATGAACATCAGCTAGCAAAAGAAAGTAACAGAGT		4440
QY	4441	GATTCCTGCTGCTATTTACTGCTAAAAA	AAAAA	4481
Db	4441	GATTCCTGCTGCTATTTACTGCTAAAAA	AAAAA	4481
RESULT 3				
AR447280				
LOCUS	AR447280	4481 bp	mRNA	linear
DEFINITION	Sequence 2 from patent US 6673535.			PAT 20-FEB-2004
ACCESSION	AR447280			
VERSION	AR447280.1	GI:42675575		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	Pulst, S.M.			
TITLE	Methods of detecting spinocerebellar Ataxia-2 nucleic acids			
JOURNAL	Patent: US 6673535-A 2 06-JAN-2004;			
	Cedars-Sinai Medical Center, Los Angeles, CA			
FEATURES	Location/Qualifiers			
source	1. .4481			
	/organism="unknown"			
	/mol_type="mRNA"			
ORIGIN				
Query Match 100.0%; Score 4481; DB 6; Length 4481;				
Best Local Similarity 100.0%; Pred. No. 0;				
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ACCCCGAGAAAGCAACCCAGCGCGCGCTCTCACGTTGCCCTCCCGGCCCGGG	60	
Db	1	ACCCCGAGAAAGCAACCCAGCGCGCGCTCTCACGTTGCCCTCCCGGCCCGGG	60	
QY	61	GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG	120	
Db	61	GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG	120	
QY	121	CACCTCCGCTCCACCCGCGGCTCGGCGCGCGCTCCGATCGCTCAGCGGCGCA	180	
Db	121	CACCTCCGCTCCACCCGCGGCTCGGCGCGCGCTCCGATCGCTCAGCGGCGCA	180	
QY	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCCAGGTGGCC	240	
Db	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCCAGGTGGCC	240	
QY	241	GGGTGGCGCTCGCTCCAGCGCGCGCGAGCGGGCGGGTGGCGCGCC	300	
Db	241	GGGTGGCGCTCGCTCCAGCGCGCGCGAGCGGGCGGGTGGCGCGCC	300	
QY	301	CGGGACCGTATCCCTTCGCGCGCCCTCCCGCGCGGCCCGGCCCTCCCTCCCGG	360	

Db	301	CGGGACCGTATCCCTTCCGCGCGCCCTCCCGCGCGGCCCTCCCTCCCTCCCGG	360	
QY	361	CAGAGCTCGCCTCCCTCCGCTCAGACTGTTTGTAGCAACGGCGGGCGCGC	420	
Db	361	CAGAGCTCGCCTCCCTCCGCTCAGACTGTTTGTAGCAACGGCGGGCGCGC	420	
QY	421	TTTGGGCCCGGCTCCGGCGGCTCCTTGGTCTCGGGGGGCTCCCGGCCCTTGTGTCG	480	
Db	421	TTTGGGCCCGGCTCCGGCGGCTCCTTGGTCTCGGGGGGCTCCCGGCCCTTGTGTCG	480	
QY	481	GTCTCTCTCCCTCGCAGCCCGGGCGCCCTCCGGCGCGCAACCCGGCGCTCCCGG	540	
Db	481	GTCTCTCTCCCTCGCAGCCCGGGCGCCCTCCGGCGCGCAACCCGGCGCTCCCGG	540	
QY	541	CTCGCGCCCGTGGGTCCCGCGCGCTCTCTTGGGGCGCGCGGCTCCCGG	600	
Db	541	CTCGCGCCCGTGGGTCCCGCGCGCTCTCTTGGGGCGCGCGGCTCCCGG	600	
QY	601	TGTCCTCCCGCGGCTGGAGCGGTGTATGGGCCCTTACCATGTCTGCTGAAGCCCG	660	
Db	601	TGTCCTCCCGCGGCTGGAGCGGTGTATGGGCCCTTACCATGTCTGCTGAAGCCCG	660	
QY	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720	
Db	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720	
QY	721	CAGCGCGCGCGGCTGCCAATGTCCGAAGCCCGCGCGCAGCGGCTTCTAGCGTCG	780	
Db	721	CAGCGCGCGCGGCTGCCAATGTCCGAAGCCCGCGCGCAGCGGCTTCTAGCGTCG	780	
QY	781	CCCGCGCGCGCTTCCGCTCTCGTCTCTCGTCTCTCGTCTCTCGGCTCCCG	840	
Db	781	CCCGCGCGCGCTTCCGCTCTCGTCTCTCGTCTCTCGTCTCTCGGCTCCCG	840	
QY	841	TCCTCGGTGTCGCGGACCTCCGCGGGGAGGCCCGGCTTGGCAGAGTTCGAAC	900	
Db	841	TCCTCGGTGTCGCGGACCTCCGCGGGGAGGCCCGGCTTGGCAGAGTTCGAAC	900	
QY	901	AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGTATGGAATCTATGCAAA	960	
Db	901	AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGTATGGAATCTATGCAAA	960	
QY	961	ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAA	1020	
Db	961	ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAA	1020	
QY	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAGTGTGATTTGGTACTTTGAT	1080	
Db	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAGTGTGATTTGGTACTTTGAT	1080	
QY	1081	GCATATGAGAAAGTACAGAAATCCAGTTCCGGGGCGGAAACCGTGAAGAAAT	1140	
Db	1081	GCATATGAGAAAGTACAGAAATCCAGTTCCGGGGCGGAAACCGTGAAGAAAT	1140	
QY	1141	ATTTGTTTCAAATGTTTACAGCTTTGTTGTGTACAGTTTAAAGATATGCAAT	1200	
Db	1141	ATTTGTTTCAAATGTTTACAGCTTTGTTGTGTACAGTTTAAAGATATGCAAT	1200	
QY	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTAAAGTAAAG	1260	
Db	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTAAAGTAAAG	1260	
QY	1261	GAGAAGACCTGGAGCCCTGGGATGCGAGTGAATCAGAGCAATAGGAACTT	1320	
Db	1261	GAGAAGACCTGGAGCCCTGGGATGCGAGTGAATCAGAGCAATAGGAACTT	1320	
QY	1321	TTGAAATATGAGTATCTTAATGGATGGATCCCAATGATATGTTTCGATATA	1380	
Db	1321	TTGAAATATGAGTATCTTAATGGATGGATCCCAATGATATGTTTCGATATA	1380	
QY	1381	AAATATGTTGTTGTTTACGTTATGATGAGTATCTTCTCGTATACAGTCCCT	1440	

Db 1381 AATTATGGTGTAGTGTCTACGTAATGATAGCAGTTTATCTTCGTATACAGTGCCCTTAGAA 1440  
Qy 1441 AGAGATAACTCAGAGAAATTTTTTAAACCGGGAAGCAAGGGGCAAAACGAGTTAGCAGAGAA 1500  
Db 1441 AGAGATAACTCAGAGAAATTTTTTAAACCGGGAAGCAAGGGGCAAAACGAGTTAGCAGAGAA 1500  
Qy 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATAGGAGTGTAG 1560  
Db 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATAGGAGTGTAG 1560  
Qy 1561 GAAAGAAAATACACAGCAGTTCAGAGAAATCCAGTGAACGTGAGGGGCAACAGCATAAAC 1620  
Db 1561 GAAAGAAAATACACAGCAGTTCAGAGAAATCCAGTGAACGTGAGGGGCAACAGCATAAAC 1620  
Qy 1621 ACTAGGGAAAATPAATATATATCTCTCGCAAAAGAAATPAGAGAGTCAATATCTCTGGGGA 1680  
Db 1621 ACTAGGGAAAATPAATATATATCTCTCGCAAAAGAAATPAGAGAGTCAATATCTCTGGGGA 1680  
Qy 1681 AGTGGGAGACAGAAATTCACCGGTATGGCCAGCCCTGGATCGGGCTCCATGCCATCAAGA 1740  
Db 1681 AGTGGGAGACAGAAATTCACCGGTATGGCCAGCCCTGGATCGGGCTCCATGCCATCAAGA 1740  
Qy 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATTCGTGTTCAAGCAAAAGAGTAGTTAAT 1800  
Db 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATTCGTGTTCAAGCAAAAGAGTAGTTAAT 1800  
Qy 1801 GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTCGCCCACTTCTCGCTAC 1860  
Db 1801 GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTCGCCCACTTCTCGCTAC 1860  
Qy 1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGGCAGCACCCCTACACGGCGGCCCTCCAGG 1920  
Db 1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGGCAGCACCCCTACACGGCGGCCCTCCAGG 1920  
Qy 1921 CCCCCCTCGCGGCATTCAGAGCCCGCTCTACCCCTCTGCTCATGGTTCTCGAGTCTCT 1980  
Db 1921 CCCCCCTCGCGGCATTCAGAGCCCGCTCTACCCCTCTGCTCATGGTTCTCGAGTCTCT 1980  
Qy 1981 GTCTCTACTATGCTTAAGCAGTGTCTTCAGAGGGCCCTCCAAGGATGTCCTCCAAAGGCC 2040  
Db 1981 GTCTCTACTATGCTTAAGCAGTGTCTTCAGAGGGCCCTCCAAGGATGTCCTCCAAAGGCC 2040  
Qy 2041 CAGCGACATCTCGAAATCAGAGTTTCTGCTGGGAGGGTTTCCATATCCAGTGGCCCTA 2100  
Db 2041 CAGCGACATCTCGAAATCAGAGTTTCTGCTGGGAGGGTTTCCATATCCAGTGGCCCTA 2100  
Qy 2101 GAATTTGTATCCACAACCCACCCAGTGAAGCAGTCTCTCCAGTAGCAAGGACCACT 2160  
Db 2101 GAATTTGTATCCACAACCCACCCAGTGAAGCAGTCTCTCCAGTAGCAAGGACCACT 2160  
Qy 2161 CCCTCGGGGGGAACGTGGTCAATCAGTGGTCAAGTGGGTTCCAAGATTAATCCCTTAAACT 2220  
Db 2161 CCCTCGGGGGGAACGTGGTCAATCAGTGGTCAAGTGGGTTCCAAGATTAATCCCTTAAACT 2220  
Qy 2221 CATAGACCGAGTCTCCAGACAGAGATTTGGAAATACCCAGTGGGCGCAGTTCTT 2280  
Db 2221 CATAGACCGAGTCTCCAGACAGAGATTTGGAAATACCCAGTGGGCGCAGTTCTT 2280  
Qy 2281 GCTTCTCCCAAGCTGTATTTTCAACTGAAGCTGTGGCATGCTTATCCAGCTGCA 2340  
Db 2281 GCTTCTCCCAAGCTGTATTTTCAACTGAAGCTGTGGCATGCTTATCCAGCTGCA 2340  
Qy 2341 TCTCTACGCTCTAGTCTCGATCGAAACAGAGCTGTATACCCCTTCTAGTGGGCTAAA 2400  
Db 2341 TCTCTACGCTCTAGTCTCGATCGAAACAGAGCTGTATACCCCTTCTAGTGGGCTAAA 2400  
Qy 2401 GATTCAGGGTTCAGAGTACAGGGAGAACTCTCTCGAGGGAAATAAGAAATATATAA 2460  
Db 2401 GATTCAGGGTTCAGAGTACAGGGAGAACTCTCTCGAGGGAAATAAGAAATATATAA 2460  
Qy 2461 CCAATGAACATCAGTCTCTCAAGCTGAAACAAAGGTATATACCCAGTTGT 2520  
Db 2461 CCAATGAACATCAGTCTCTCAAGCTGAAACAAAGGTATATACCCAGTTGT 2520

Qy 2521 TCTGAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGGTTACAG 2580  
Db 2521 TCTGAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGGTTACAG 2580  
Qy 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAAACAATAAGAGAGGAGAAAA 2640  
Db 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAAACAATAAGAGAGGAGAAAA 2640  
Qy 2641 TCAAGAGATTTGATCAAAGACAAATTTGAACCAAGTGTAAAGGATTTCTTCAATTCAAAAT 2700  
Db 2641 TCAAGAGATTTGATCAAAGACAAATTTGAACCAAGTGTAAAGGATTTCTTCAATTCAAAAT 2700  
Qy 2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760  
Db 2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760  
Qy 2761 ATACTTAGTAACACGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCCAGACT 2820  
Db 2761 ATACTTAGTAACACGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCCAGACT 2820  
Qy 2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAGAGAAAGACGACCTGAG 2880  
Db 2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAGAGAAAGACGACCTGAG 2880  
Qy 2881 CAAGTTAGGAATCAACATTGAATCCCAATGCAAGAGGATTCACCCAGTTCTTCTCT 2940  
Db 2881 CAAGTTAGGAATCAACATTGAATCCCAATGCAAGAGGATTCACCCAGTTCTTCTCT 2940  
Qy 2941 CAGCCAAAGCCCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCCATCT 3000  
Db 2941 CAGCCAAAGCCCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCCATCT 3000  
Qy 3001 ATGGTGGTCTATCAACAGCCAACTCCAGTTTATATCTAGCCTGTTTGTGTTGACCAAAAT 3060  
Db 3001 ATGGTGGTCTATCAACAGCCAACTCCAGTTTATATCTAGCCTGTTTGTGTTGACCAAAAT 3060  
Qy 3061 ATGATGTATCCAGTCCCGAGTGGCCAGCGTGCACCTTTATACCCAATACCTATGAGC 3120  
Db 3061 ATGATGTATCCAGTCCCGAGTGGCCAGCGTGCACCTTTATACCCAATACCTATGAGC 3120  
Qy 3121 CCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCAACAGCG 3180  
Db 3121 CCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCAACAGCG 3180  
Qy 3181 CAAAGACAGCATCATCAGATGTCATGTGCAACCCAGCGTCAAGCGGGGCCACCGATT 3240  
Db 3181 CAAAGACAGCATCATCAGATGTCATGTGCAACCCAGCGTCAAGCGGGGCCACCGATT 3240  
Qy 3241 GCAGCCACCCACAGCTTACTCCAGCGCAATATGTTGCTACAGTCTCAGCAGTTCCTCA 3300  
Db 3241 GCAGCCACCCACAGCTTACTCCAGCGCAATATGTTGCTACAGTCTCAGCAGTTCCTCA 3300  
Qy 3301 AATCAGCCCTGTTTTCAGCATGTGCACATATCAGTCTCAGCATCTCTCATGTATAGT 3360  
Db 3301 AATCAGCCCTGTTTTCAGCATGTGCACATATCAGTCTCAGCATCTCTCATGTATAGT 3360  
Qy 3361 CCTGTAATACAGGGTAATGCTAGAAATGAGCCACCAACACAGCCCGCAGCTTGTGTTA 3420  
Db 3361 CCTGTAATACAGGGTAATGCTAGAAATGAGCCACCAACACAGCCCGCAGCTTGTGTTA 3420  
Qy 3421 GTATCTTCTTACGAACTCAGTACGGGCTCATGAGCAGACGCATGCGATGTATGATGT 3480  
Db 3421 GTATCTTCTTACGAACTCAGTACGGGCTCATGAGCAGACGCATGCGATGTATGATGT 3480  
Qy 3481 CCCAAATACCATACAAAGAGAGCAAGCCCTCTTTCTACTTTGCGCAATTTCCACGGGC 3540  
Db 3481 CCCAAATACCATACAAAGAGAGCAAGCCCTCTTTCTACTTTGCGCAATTTCCACGGGC 3540  
Qy 3541 TCCCTTGCTCAGCAGTATGCGCACCTCAACGCTACCTGACCCACATACTCCACACCT 3600  
Db 3541 TCCCTTGCTCAGCAGTATGCGCACCTCAACGCTACCTGACCCACATACTCCACACCT 3600



QY 901 AGTAACAAAGGAGTGCCTCAGTCTACGATTTCTTTTGATGGAATCTATGCCAAATATGAGG 960  
DB 901 AGTAACAAAGGAGTGCCTCAGTCTACGATTTCTTTTGATGGAATCTATGCCAAATATGAGG 960  
QY 961 ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGGA 1020  
DB 961 ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGGA 1020  
QY 1021 GGTATATATGAAGGAGTTTTTAAACTTACAGTCCGAACTGTGATTTGGTACTTGTATGCC 1080  
DB 1021 GGTATATATGAAGGAGTTTTTAAACTTACAGTCCGAACTGTGATTTGGTACTTGTATGCC 1080  
QY 1081 GCACATGAGAAAGATACAGAAATCCAGTTCGGGGCGGAAACGTGTGAAGAAATATGAGAGT 1140  
DB 1081 GCACATGAGAAAGATACAGAAATCCAGTTCGGGGCGGAAACGTGTGAAGAAATATGAGAGT 1140  
QY 1141 ATTTTGTTCAAATGTTTCAGACTTTGTTGTGTGATACAGTTTAAAGATATGAGACTCCAGTTAT 1200  
DB 1141 ATTTTGTTCAAATGTTTCAGACTTTGTTGTGTGATACAGTTTAAAGATATGAGACTCCAGTTAT 1200  
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGCTAAAGTGAATGGCGAAACACAAA 1260  
DB 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGCTAAAGTGAATGGCGAAACACAAA 1260  
QY 1261 GAGAAAGACTGGAGCCCTGGGATGAGGTGAACCTCACAGCCAAATGAGGAACTTGAGGCT 1320  
DB 1261 GAGAAAGACTGGAGCCCTGGGATGAGGTGAACCTCACAGCCAAATGAGGAACTTGAGGCT 1320  
QY 1321 TTGGAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAGAA 1380  
DB 1321 TTGGAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAGAA 1380  
QY 1381 AATTATGTTGATGTTCTAGTATGATAGCAGTTTATCTTCGTATACAGTGGCCCTTAGAA 1440  
DB 1381 AATTATGTTGATGTTCTAGTATGATAGCAGTTTATCTTCGTATACAGTGGCCCTTAGAA 1440  
QY 1441 AGAGATAACTCAGAGAAATTTTTTAAACCGGGAAGCAAGGCGCAACCAAGTTAGCAGAA 1500  
DB 1441 AGAGATAACTCAGAGAAATTTTTTAAACCGGGAAGCAAGGCGCAACCAAGTTAGCAGAA 1500  
QY 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATCATATAGGAGTGA 1560  
DB 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATCATATAGGAGTGA 1560  
QY 1561 GAAGAAAAATACACAGCAGTTTACAGAAATTTCCAGTGAACGTGAGGGCGACACATATAAC 1620  
DB 1561 GAAGAAAAATACACAGCAGTTTACAGAAATTTCCAGTGAACGTGAGGGCGACACATATAAC 1620  
QY 1621 ACTAGGAAAAATAAATATATTCCTCCTGGGCAAGAAATAGAGAGTCTATCTCGGGGA 1680  
DB 1621 ACTAGGAAAAATAAATATATTCCTCCTGGGCAAGAAATAGAGAGTCTATCTCGGGGA 1680  
QY 1681 AGTGGAGACAGAAATTCACCGGTATGGGCAGCCTGGATCGGGCTCCATGCCATCAAGA 1740  
DB 1681 AGTGGAGACAGAAATTCACCGGTATGGGCAGCCTGGATCGGGCTCCATGCCATCAAGA 1740  
QY 1741 TCCACTTCTCACACTTCAGATTTCAACCGAAATTCCTGGTTTCAGACCAAGAGTAGTTAAT 1800  
DB 1741 TCCACTTCTCACACTTCAGATTTCAACCGAAATTCCTGGTTTCAGACCAAGAGTAGTTAAT 1800  
QY 1801 GGAGGTGTTCCCTGGCCATCGCCTTGCCTATCTCTCTCTCGCCACCTTCTCGCTAC 1860  
DB 1801 GGAGGTGTTCCCTGGCCATCGCCTTGCCTATCTCTCTCTCGCCACCTTCTCGCTAC 1860  
QY 1861 CAGTCAGGTCCCAACTCTCTTTCACCTTCGGGCGACCCCTACACGGCGCCCTCCAGG 1920  
DB 1861 CAGTCAGGTCCCAACTCTCTTTCACCTTCGGGCGACCCCTACACGGCGCCCTCCAGG 1920  
QY 1921 CCCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGCTCCT 1980  
DB 1921 CCCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGCTCCT 1980

QY 1981 GTCTCTACTATGCTTAAACGATGTCTTTCAGAAAGGCGCTTCAAGGATGTCCCAAGGCC 2040  
DB 1981 GTCTCTACTATGCTTAAACGATGTCTTTCAGAAAGGCGCTTCAAGGATGTCCCAAGGCC 2040  
QY 2041 CAGCGACATCTCTCGAAATCAGAGAGTTTCTGCTGGAGGGGTTTCCATATCCAGTGGCCTA 2100  
DB 2041 CAGCGACATCTCTCGAAATCAGAGAGTTTCTGCTGGAGGGGTTTCCATATCCAGTGGCCTA 2100  
QY 2101 GAAATTTGTTATCCCAACCCAGTGAAGCAGTACTCTCCAGTAGCAGGACCACT 2160  
DB 2101 GAAATTTGTTATCCCAACCCAGTGAAGCAGTACTCTCCAGTAGCAGGACCACT 2160  
QY 2161 CCCTCGGGGGGAACTGTGTGATCAGTGTGAGTGGGTTTCAAGATTTATCCCTTAAACT 2220  
DB 2161 CCCTCGGGGGGAACTGTGTGATCAGTGTGAGTGGGTTTCAAGATTTATCCCTTAAACT 2220  
QY 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGAAATATACCCCGAGTGGGCGAGTTCT 2280  
DB 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGAAATATACCCCGAGTGGGCGAGTTCT 2280  
QY 2281 GCTTCTCCCAAGCTGTTATTTCCAACTGAAGCTGTTGCCATGCTATTTCCAGTGA 2340  
DB 2281 GCTTCTCCCAAGCTGTTATTTCCAACTGAAGCTGTTGCCATGCTATTTCCAGTGA 2340  
QY 2341 TCTCTAGCGCTGTAGTCTGCATCGAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA 2400  
DB 2341 TCTCTAGCGCTGTAGTCTGCATCGAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA 2400  
QY 2401 GATTTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAGAAATATTTAA 2460  
DB 2401 GATTTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAGAAATATTTAA 2460  
QY 2461 CCCAATGAACATCACCTAGTCTTCTCAAAAGCTGAAAAAAGGATATATCAACAGTTGT 2520  
DB 2461 CCCAATGAACATCACCTAGTCTTCTCAAAAGCTGAAAAAAGGATATATCAACAGTTGT 2520  
QY 2521 TCTGAACATAGAAACAGATTTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGTTACAG 2580  
DB 2521 TCTGAACATAGAAACAGATTTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGTTACAG 2580  
QY 2581 CCAAGTTCTACTTCTGAACTTATGATCAACTACTTAAACAAAAATAGAGAGGAGAAAAA 2640  
DB 2581 CCAAGTTCTACTTCTGAACTTATGATCAACTACTTAAACAAAAATAGAGAGGAGAAAAA 2640  
QY 2641 TCAAGAGATTTGATCAAAAGACAAAAATTGAACCAAGTGTCTTCAATTTTCAATTTGAAAAT 2700  
DB 2641 TCAAGAGATTTGATCAAAAGACAAAAATTGAACCAAGTGTCTTCAATTTTCAATTTGAAAAT 2700  
QY 2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCGAGCATTTTCCCTTCA 2760  
DB 2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCGAGCATTTTCCCTTCA 2760  
QY 2761 ATACTTAGTTAAACAGGAGCACAGAGGGGACCTGAGTCACTTCCCAAGGGTTTACAGACT 2820  
DB 2761 ATACTTAGTTAAACAGGAGCACAGAGGGGACCTGAGTCACTTCCCAAGGGTTTACAGACT 2820  
QY 2821 TCCAGCCAGCATGTAAAACAGAGAAAGACGATAGAGAAAGAGAAAGACGAGCTGAG 2880  
DB 2821 TCCAGCCAGCATGTAAAACAGAGAAAGACGATAGAGAAAGAGAAAGACGAGCTGAG 2880  
QY 2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCTCT 2940  
DB 2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCTCT 2940  
QY 2941 CAGCCAAAGCCTTCTACTACCCCACTTCCCTCGGCTTCAAGCAAACTAGCCCATCT 3000  
DB 2941 CAGCCAAAGCCTTCTACTACCCCACTTCCCTCGGCTTCAAGCAAACTAGCCCATCT 3000  
QY 3001 ATGGTGGGTATCAACAGCAAACTCCAGTTTATCTCAGCCTGTTTGTGTTTGCACCAAT 3060  
DB 3001 ATGGTGGGTATCAACAGCAAACTCCAGTTTATCTCAGCCTGTTTGTGTTTGCACCAAT 3060  
QY 3061 ATGATGATTCAGTCCAGTGAGCCCGGCTGCAACCTTTTATACCCAAATACCTATGACG 3120

||||| 3061 ATGATGTATCAGTCCCGAGTGGCCGCGTGAACCTTTATACCCAAATACCTATGACG 3120  
||||| 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCCAACAGCGG 3180  
||||| 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCCAACAGCGG 3180  
||||| 3181 CAAGACCAGCATCATCAGATGCCATGATGACCCAGCGGTGAGCAGCGGGCCACCGATT 3240  
||||| 3181 CAAGACCAGCATCATCAGATGCCATGATGACCCAGCGGTGAGCAGCGGGCCACCGATT 3240  
||||| 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTACAGTCCCTCAGCAGTTCCCA 3300  
||||| 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTACAGTCCCTCAGCAGTTCCCA 3300  
||||| 3301 AATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCCCTCATGTCTATAGT 3360  
||||| 3301 AATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCCCTCATGTCTATAGT 3360  
||||| 3361 CCTGTAAATACAGGGTAATGCTAGAAATGATGGACCAACACACAGCGCCAGCTGGTTTA 3420  
||||| 3361 CCTGTAAATACAGGGTAATGCTAGAAATGATGGACCAACACACAGCGCCAGCTGGTTTA 3420  
||||| 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCGCATGCGATGTATGATGT 3480  
||||| 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCGCATGCGATGTATGATGT 3480  
||||| 3481 CCCAAATFACCATCAACAAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTTCCACGGGC 3540  
||||| 3481 CCCAAATFACCATCAACAAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTTCCACGGGC 3540  
||||| 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTGTCACCCACCATACTCCACACCT 3600  
||||| 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTGTCACCCACCATACTCCACACCT 3600  
||||| 3601 CAGCCTTCAGTACCCCACTGGACAGCAGCAAGCCAAACATGTTGGAAGTATCTCTGCA 3660  
||||| 3601 CAGCCTTCAGTACCCCACTGGACAGCAGCAAGCCAAACATGTTGGAAGTATCTCTGCA 3660  
||||| 3661 CCCAGTCTGTTACACCATCAGCACCGCGCCGCTTCCATCTCGGCCAGTGCA 3720  
||||| 3661 CCCAGTCTGTTACACCATCAGCACCGCGCCGCTTCCATCTCGGCCAGTGCA 3720  
||||| 3721 CAGCAGCAGTACGCAATTTACACGCGGGCTTGGCCCAACTCCACCTCCATGACACCT 3780  
||||| 3721 CAGCAGCAGTACGCAATTTACACGCGGGCTTGGCCCAACTCCACCTCCATGACACCT 3780  
||||| 3781 GCCTCCACAGCAGTCGCCACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
||||| 3781 GCCTCCACAGCAGTCGCCACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
||||| 3841 ATCCATCTTCTCAGTTCAGCGCGGTATACCAACCCACCCACATGGCCAGCTACCT 3900  
||||| 3841 ATCCATCTTCTCAGTTCAGCGCGGTATACCAACCCACCCACATGGCCAGCTACCT 3900  
||||| 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTTCTCATCAACTGCCCATCGCCCAATGATG 3960  
||||| 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTTCTCATCAACTGCCCATCGCCCAATGATG 3960  
||||| 4021 CTAATGACACACAGCAGCAGCGGGTCCCGAGCGCCCTCGCTCAAAAGTGCACTACAG 4020  
||||| 3961 CTAATGACACACAGCAGCAGCGGGTCCCGAGCGCCCTCGCTCAAAAGTGCACTACAG 4020  
||||| 4021 CCCATTCCAGTCTGCACACAGCGCATTTCCCTCATATAGCGCACCTTTCAGTCAAGCC 4080  
||||| 4021 CCCATTCCAGTCTGCACACAGCGCATTTCCCTCATATAGCGCACCTTTCAGTCAAGCC 4080  
||||| 4081 CACCACCAACAGCAGTGTGAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCCTC 4140  
||||| 4081 CACCACCAACAGCAGTGTGAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCCTC 4140  
||||| 4141 CCTTCTACTGCTTACCAACTGGAGGACACAGAACTAGAAATTTTATTTTGT 4200

Db 4141 CCTTCTACTGCTTACTCAACTGGAAGCAGAGAAAACCTAGAAATTTTATTTTGT 4200  
Qy 4201 TAAATATATATGTTGATTTCTTCTTAACATCCAAATAGGAATGCTAAACAGTTTCACTTGCAG 4260  
Db 4201 TAAATATATATGTTGATTTCTTCTTAACATCCAAATAGGAATGCTAAACAGTTTCACTTGCAG 4260  
Qy 4261 TGAAGATATCTTGGACCGAGTAGAGGCAATTTAGGAATCTTGGGGCTATTTCCATTAATTTCCA 4320  
Db 4261 TGAAGATATCTTGGACCGAGTAGAGGCAATTTAGGAATCTTGGGGCTATTTCCATTAATTTCCA 4320  
Qy 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGCAAGTTATTT 4380  
Db 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGCAAGTTATTT 4380  
Qy 4381 ATTTTAAATTAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAAGAAAGTAAACAAGAGT 4440  
Db 4381 ATTTTAAATTAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAAGAAAGTAAACAAGAGT 4440  
Qy 4441 GATTTCTGCTGCTATTAATGCTCTAAAAAATAAAAAA 4481  
Db 4441 GATTTCTGCTGCTATTAATGCTCTAAAAAATAAAAAA 4481

RESULT 5  
AX392465  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

AX392465  
Sequence 1 from Patent WO0216417.  
AX392465  
AX392465.1 GI:19700764  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1  
Pulst,S.M. and Huynh,D.P.  
transgenic animal model for neurodegenerative disease and uses  
thereof  
Patent: WO 0216417-A 1 28-FEB-2002;  
CEDARS-SINAI MEDICAL CENTER (US)  
Location/Qualifiers  
1. .4481  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
163..4101  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD29012.1"  
/db\_xref="GI:19700765"  
/translation="MRSAAAPRSPAVATSRRRFAAARWPGWRSLORPARSRGRGGG  
AAGPYPSAAPPFGPPGPPSPRSQSSASDCPCSGNGNGGAFRPGSRLLGLGGPPR  
PFVVLLPLASPAPPAPATRASPLGARASPPRGSVSLARPACCPACBPVYGLPT  
MSLQPOQQQQQQQQQQQQQQPPAAANVRKPGSGLLASPAAPSPSSSV  
SSCATAPSVVAATSGGRLGRLGRNSKGLPQSTISFDGIYANMRMHILTSVVG  
SKSRVQKNGGIYBGVFTYSPKCDLVLDAAHEKSTESSGPKKEEIMESILFKCSDP  
VVQFKMDMSYAKRDAFTDSAI SAKVNGEHKEDLEWDAGEI.TANEELALENDVS  
NGWPDNDMRYNEENYGVSTSDSSLSYTVPLERDSEBFLKREARANQLAEIES  
AQYKARVALENDDRSEBEKYAVORNSSREGEHSINTRENKIIPPGQRNREVI SWGSG  
RQNSPRLPGQSGSMPSRSTSHTSDFNPSGSDQRVYNGVWPSPCPSPSPSPSR  
QSGNSLPPRAATPTPPSPRPPSPRPPSPRPPSPRPPSPRPPSPRPPSPRPPSR  
KAORHPNRVSRAGSGISGGLBFSVSHNPPEAATPPVARTSPSGTWSVSVGVPLI  
SPKTHRPSPRONSIGNTPSGPLVASPOAGIIPTEAVAMP I PAASPPTPASPNRAVT  
PSSAKDSRLQDQRNSPAGNKENIKENETSPSPSKAENKGISPVSEHRKQIDDLKX  
FKNDFRLQPSSTSEMDLLNKRGEKSRDLIKDIEPSAKDSFIENSSNCRSTLN  
KNPSPIPSILSNTEHKEPVEVTSQGVTSPPACKQEKDDKEEKDAAEQVRKSTLN  
PNAKEFNPRSPQPKSTPTSPRQAQSPSPSMVHQQTPTVYTPVCFAFNMYVPV  
SPQGVQLYPTPMTFVNQAKTVRVPNMPQORQDQHQSHSMHMPASAGPPAATP  
PAYSTQVAVSPQGFNPQVHPVYQSHPHVYSPVIOGNARMPAPPTHAOPGLVS  
SSATQYAGBQTHMYACPKLPYNKETSPSFYFAISTGSLAQOYAHNPATLHPHTPH  
QPSATPTGQQQSGHGHAPSPVQHQHQHAQALHLASPOQSAIYHAGLAPTPPSM  
TPASNTSPQNSFPAAQQTVTTHPSHVQPAYTNPPHMAHVQAHVQSGMSPFHTAH



ORIGIN									
APMMLMTTQPPGGPQALAQSAQLQIPVSTTAHPYMTHPSPVQAHQOOL"									
Query Match 100.0%; Score 4481; DB 6; Length 4481;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ACCCCGAGAAAGCAACCCAGCGCGCCCGCTCCTCAGGTGTCCTCCCGCGCCCGGG	60						
Db	1	ACCCCGAGAAAGCAACCCAGCGCGCCCGCTCCTCAGGTGTCCTCCCGCGCCCGGG	60						
Qy	61	GCCACCTCAGCTTCGCTTCGGTCTGACCCCTCGACTTCGGTAAAGAGTCCCTATCCG	120						
Db	61	GCCACCTCAGCTTCGCTTCGGTCTGACCCCTCGACTTCGGTAAAGAGTCCCTATCCG	120						
Qy	121	CACCTTCGCTCCACCCCGCGCTCGGCGCGCCCTCCGATGCGTCCAGCGCCGCA	180						
Db	121	CACCTTCGCTCCACCCCGCGCTCGGCGCGCCCTCCGATGCGTCCAGCGCCGCA	180						
Qy	181	GCTCTCGAGTCCCGCGGTGCGACCGAGTCTCGCGCTTCGCGTAAAGAGTCCCTATCCG	240						
Db	181	GCTCTCGAGTCCCGCGGTGCGACCGAGTCTCGCGCTTCGCGTAAAGAGTCCCTATCCG	240						
Qy	241	GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	300						
Db	241	GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	300						
Qy	301	CGGAGACGTATCCCTCCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCG	360						
Db	301	CGGAGACGTATCCCTCCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCG	360						
Qy	361	CAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAACCGCGCGCGCGCG	420						
Db	361	CAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAACCGCGCGCGCGCG	420						
Qy	421	TTTCCGCGCGCTCCCGCGCGCTTCGTCTCGCGCGCGCTCCCGCGCGCTTCGTCTC	480						
Db	421	TTTCCGCGCGCTCCCGCGCGCTTCGTCTCGCGCGCGCTCCCGCGCGCTTCGTCTC	480						
Qy	481	GTCTTCTTCCCTCCCGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCTCCCGCG	540						
Db	481	GTCTTCTTCCCTCCCGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCTCCCGCG	540						
Qy	541	CTCGCGCGCGCTCCCGCGCGCTTCGTCTCGCGCGCGCTTCGTCTCGCGCGCG	600						
Db	541	CTCGCGCGCGCTCCCGCGCGCTTCGTCTCGCGCGCGCTTCGTCTCGCGCGCG	600						
Qy	601	TGTCTCCCGCGCTTCGCGCGCTTCGTCTCGCGCGCGCTTCGTCTCGCGCGCG	660						
Db	601	TGTCTCCCGCGCTTCGCGCGCTTCGTCTCGCGCGCGCTTCGTCTCGCGCGCG	660						
Qy	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720						
Db	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720						
Qy	721	CAGCGCGCGCGCTCCGCTCAGTTCGCAAGCCCGCGCGCGCGCGCGCGCGCG	780						
Db	721	CAGCGCGCGCGCTCCGCTCAGTTCGCAAGCCCGCGCGCGCGCGCGCGCGCG	780						
Qy	781	CCGCGCGCGCGCTTCGCGCGCTTCGTCTCGCGCGCGCTTCGTCTCGCGCGCG	840						
Db	781	CCGCGCGCGCGCTTCGCGCGCTTCGTCTCGCGCGCGCTTCGTCTCGCGCGCG	840						
Qy	841	TCCTCGGTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCG	900						
Db	841	TCCTCGGTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCGCTTCGCGCG	900						
Qy	901	AGTAAACAAAGGACTGCTCAGTTCAGTTCGTAATTTTGTGATGGAATCTATGGA	960						
Db	901	AGTAAACAAAGGACTGCTCAGTTCAGTTCGTAATTTTGTGATGGAATCTATGGA	960						
Qy	961	ATGGTTTCATATCTTACATCAGTTGTTGGTCCCAATGTGAAGTACAGTGAATAATG	1020						
Db	961	ATGGTTTCATATCTTACATCAGTTGTTGGTCCCAATGTGAAGTACAGTGAATAATG	1020						

Db	961	ATGGTTTCATATCTTACATCAGTTGTTGGTCCCAATGTGAAGTACAGTGAATAATG	1020
Qy	1021	GGTATATATGAAGGAGTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC	1080
Db	1021	GGTATATATGAAGGAGTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC	1080
Qy	1081	GCACATGAGAAAGTACAGATCCAGTTCGGGGCGGAACCTGAAGAAATATCGAGAGT	1140
Db	1081	GCACATGAGAAAGTACAGAAATCAGTTCGGGGCGGAACCTGAAGAAATATCGAGAGT	1140
Qy	1141	ATTTTGTTCAAATGTTTACAGCTTTGTTGTGTACAGTTTAAAGATATGAGCTCCAGTTAT	1200
Db	1141	ATTTTGTTCAAATGTTTACAGCTTTGTTGTGTACAGTTTAAAGATATGAGCTCCAGTTAT	1200
Qy	1201	GCAAAAAGAGATGCTTTTACCTGCTATCAGTCTAAAGTGAATGAGGCAACACAAA	1260
Db	1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGAGGCAACACAAA	1260
Qy	1261	GAGAGGACCTGGAGCCTGGGATGCGAGTCAACTCAGCAATGAGGAATTTGAGGCT	1320
Db	1261	GAGAGGACCTGGAGCCTGGGATGCGAGTCAACTCAGCAATGAGGAATTTGAGGCT	1320
Qy	1321	TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAATGAAGA	1380
Db	1321	TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAATGAAGA	1380
Qy	1381	AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA	1440
Db	1381	AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA	1440
Qy	1441	AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGGCAACACAGTTAGCAAGAA	1500
Db	1441	AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGGCAACACAGTTAGCAAGAA	1500
Qy	1501	ATTGAGTCAAGTGGCCAGTCAAAAGCTCAGTGGCCCTGGAAAATGATAGAGAGTGA	1560
Db	1501	ATTGAGTCAAGTGGCCAGTCAAAAGCTCAGTGGCCCTGGAAAATGATAGAGAGTGA	1560
Qy	1561	GAAGAAAATACACAGCAGTTTCCAGTGAATTTCCAGTGAACGTGAGGGGACAGCATAAAC	1620
Db	1561	GAAGAAAATACACAGCAGTTTCCAGTGAATTTCCAGTGAACGTGAGGGGACAGCATAAAC	1620
Qy	1621	ACTAGGAAAATTAATATATCTCTCGACAAAGAAATAGAGATCATATCTCTGGGA	1680
Db	1621	ACTAGGAAAATTAATATATCTCTCGACAAAGAAATAGAGATCATATCTCTGGGA	1680
Qy	1681	AGTGGGACAGAAATTCACCGCTATGGCGCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
Db	1681	AGTGGGACAGAAATTCACCGCTATGGCGCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
Qy	1741	TCCACTTCTCACATTCAGATTTTCAACCCGAATTCGTGTTTCAGACCAAGAGTAGTTAAT	1800
Db	1741	TCCACTTCTCACATTCAGATTTTCAACCCGAATTCGTGTTTCAGACCAAGAGTAGTTAAT	1800
Qy	1801	GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTCGCCACCTCTCGCCACCTCTCGCTAC	1860
Db	1801	GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTCGCCACCTCTCGCCACCTCTCGCTAC	1860
Qy	1861	CAGTCAGTTCCTCAACTCTCTTCCACTCGGGCAGCACCCTTACACGGCGCCCTCCAG	1920
Db	1861	CAGTCAGTTCCTCAACTCTCTTCCACTCGGGCAGCACCCTTACACGGCGCCCTCCAG	1920
Qy	1921	CCCCCTCGCGCCATCCAGACCCCGCTCAACCCTCTGCTCATGTTCTCCAGTCTCT	1980
Db	1921	CCCCCTCGCGCCATCCAGACCCCGCTCAACCCTCTGCTCATGTTCTCCAGTCTCT	1980
Qy	1981	GTCTCTACTATGCTTAAAGCATGTCTTCAGAGGGCTCCCAAGTGTCCCAAGGCC	2040
Db	1981	GTCTCTACTATGCTTAAAGCATGTCTTCAGAGGGCTCCCAAGTGTCCCAAGGCC	2040
Qy	2041	CAGGCACATCTTCGAAATCAGAGTTCCTGCTGGAGGGGTTCATATCAGTGCCTTA	2100
Db	2041	CAGGCACATCTTCGAAATCAGAGTTCCTGCTGGAGGGGTTCATATCAGTGCCTTA	2100



Qy	2101	GAATTTGTATCCCAAAACCCACCCAGTGAAGAGCTACTCTCCAGTAGAAGACCAAGT	2160
Db	2101	GAATTTGTATCCCAAAACCCACCCAGTGAAGAGCTACTCTCCAGTAGAAGACCAAGT	2160
Qy	2161	CCCTCGGGGGAAAGTGGTGCATCAGTGGTCAAGTGGGTTCCAAAGTATCCCTTAAACT	2220
Db	2161	CCCTCGGGGGAAAGTGGTGCATCAGTGGTCAAGTGGGTTCCAAAGTATCCCTTAAACT	2220
Qy	2221	CATAGACCAGGCTCTCCAGACAGAACTGTTGGAATAACCCAGTGGGCCAGTTCCTT	2280
Db	2221	CATAGACCAGGCTCTCCAGACAGAACTGTTGGAATAACCCAGTGGGCCAGTTCCTT	2280
Qy	2281	GCTTCTCCCAAGCTGGTATTTATTCAACTGAAGCTGTTGCCATGCTTATCCAGCTGCA	2340
Db	2281	GCTTCTCCCAAGCTGGTATTTATTCAACTGAAGCTGTTGCCATGCTTATCCAGCTGCA	2340
Qy	2341	TCTCCTAGCCTGTAGTCTGTCATCGAACAAGAGCTGTATCCCTTCTAGTGAGGCTAAA	2400
Db	2341	TCTCCTAGCCTGTAGTCTGTCATCGAACAAGAGCTGTATCCCTTCTAGTGAGGCTAAA	2400
Qy	2401	GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCTGAGGGAATAAAGAAAAATTAAA	2460
Db	2401	GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCTGAGGGAATAAAGAAAAATTAAA	2460
Qy	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAAAAATAAGGTTGTT	2520
Db	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAAAAATAAGGTTGTT	2520
Qy	2521	TCTGAACTAGAAACAGATGATGATTTAAAGAAATTTAAGAAATGATTTAGGTTACAG	2580
Db	2521	TCTGAACTAGAAACAGATGATGATTTAAAGAAATTTAAGAAATGATTTAGGTTACAG	2580
Qy	2581	CCAAGTTCTACTTCTGAACTATGATCAACTACTGATTAACAAAAATAGAGGGAGAAAA	2640
Db	2581	CCAAGTTCTACTTCTGAACTATGATCAACTACTGATTAACAAAAATAGAGGGAGAAAA	2640
Qy	2641	TCAAGAGATTTGATCAAAAGAAATAATGAAACAAAGTGTAAAGATCTTTCAATTGAAAT	2700
Db	2641	TCAAGAGATTTGATCAAAAGAAATAATGAAACAAAGTGTAAAGATCTTTCAATTGAAAT	2700
Qy	2701	AGCAGCAGCAACTGTACAGTGGCAGCAGCGCGAATAGCCACGATTTCCCTTCA	2760
Db	2701	AGCAGCAGCAACTGTACAGTGGCAGCAGCGCGAATAGCCACGATTTCCCTTCA	2760
Qy	2761	ATACTTAGTAAACGGAGCAACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAAGCT	2820
Db	2761	ATACTTAGTAAACGGAGCAACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAAGCT	2820
Qy	2821	TCCAGCCAGCATGTAAACAAGAGAAAGACGATAGGGAAGAGAAAGACGAGCTGAG	2880
Db	2821	TCCAGCCAGCATGTAAACAAGAGAAAGACGATAGGGAAGAGAAAGACGAGCTGAG	2880
Qy	2881	CAAGTTAGGAAATCAACATTCGAATCCCAATGCAAGAGGTTCAACCCAGTTCCTTCTCT	2940
Db	2881	CAAGTTAGGAAATCAACATTCGAATCCCAATGCAAGAGGTTCAACCCAGTTCCTTCTCT	2940
Qy	2941	CAGCCAAAGCCTTCTACTACCCCAACTTCACTGGGCTCAAGCACAACCTTAGCCCATCT	3000
Db	2941	CAGCCAAAGCCTTCTACTACCCCAACTTCACTGGGCTCAAGCACAACCTTAGCCCATCT	3000
Qy	3001	ATGGTGGTCAATCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTGGCAACAAAT	3060
Db	3001	ATGGTGGTCAATCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTGGCAACAAAT	3060
Qy	3061	ATGATGTATCCAGTCCCAAGTGGCCGCTGCAACCTTTATACCCCAATACCTATGACG	3120
Db	3061	ATGATGTATCCAGTCCCAAGTGGCCGCTGCAACCTTTATACCCCAATACCTATGACG	3120
Qy	3121	CCCATGCCAGTGAATCAAGCCCAAGACATATAGACGATACCAATATGCCCCAACAGCGG	3180
Db	3121	CCCATGCCAGTGAATCAAGCCCAAGACATATAGACGATACCAATATGCCCCAACAGCGG	3180

Qy	3181	CAAGACCAGCATCATCAGATGCCATGATGCAACCAGGGTCAGCAGCGGGCCCCACCGATT	3240
Db	3181	CAAGACCAGCATCATCAGATGCCATGATGCAACCAGGGTCAGCAGCGGGCCCCACCGATT	3240
Qy	3241	GCAGCCACCCACCAGCTTACTCCACGCAATATGCTGCCTACAGTCCCTCAGCAGTCCCA	3300
Db	3241	GCAGCCACCCACCAGCTTACTCCACGCAATATGCTGCCTACAGTCCCTCAGCAGTCCCA	3300
Qy	3301	AATCAGGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT	3360
Db	3301	AATCAGGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT	3360
Qy	3361	CCTGTAATACAGGGTAATGCTAGAAATGATGGCAACCAACACAGCGCCAGCTGGTTTA	3420
Db	3361	CCTGTAATACAGGGTAATGCTAGAAATGATGGCAACCAACACAGCGCCAGCTGGTTTA	3420
Qy	3421	GTATCTTTCTTCAGCAAACTCAGTACGGGGCTCATGAGCAGAGCGCATGTATGCAATGT	3480
Db	3421	GTATCTTTCTTCAGCAAACTCAGTACGGGGCTCATGAGCAGAGCGCATGTATGCAATGT	3480
Qy	3481	CCCAAAATTCACATACAAAGGAGACAAAGCCCTTTCTTCTACTTTGGCAATTTCCACGGGC	3540
Db	3481	CCCAAAATTCACATACAAAGGAGACAAAGCCCTTTCTTCTACTTTGGCAATTTCCACGGGC	3540
Qy	3541	TCCCTTGTCTCAGCAGTATGCGCACCTTAAAGCTTACCCTGCAACCACATCTCCACACCT	3600
Db	3541	TCCCTTGTCTCAGCAGTATGCGCACCTTAAAGCTTACCCTGCAACCACATCTCCACACCT	3600
Qy	3601	CAGCCTTTCAGTACCCCTTGGACAGCAAGCAAGCAATGGTGAAGTCATCTTGCA	3660
Db	3601	CAGCCTTTCAGTACCCCTTGGACAGCAAGCAAGCAATGGTGAAGTCATCTTGCA	3660
Qy	3661	CCCAAGTCTCTTTCAGCACCCTTACCAAGCCCGCCAGGCTCTCTCATCTGGCCAGTCCA	3720
Db	3661	CCCAAGTCTCTTTCAGCACCCTTACCAAGCCCGCCAGGCTCTCTCATCTGGCCAGTCCA	3720
Qy	3721	CAGCAGAGTCAGCCATTTACCAAGCGGGGTTGCGCAACTCCACCTTCCATGACACCT	3780
Db	3721	CAGCAGAGTCAGCCATTTACCAAGCGGGGTTGCGCAACTCCACCTTCCATGACACCT	3780
Qy	3781	GCCTCCAAACAGCAGTCGCGCAAGATAGTTTTCCACAGCAGCAACAGACTGCTTTTACG	3840
Db	3781	GCCTCCAAACAGCAGTCGCGCAAGATAGTTTTCCACAGCAGCAACAGACTGCTTTTACG	3840
Qy	3841	ATCCATCTTCTTCTCAGCTTCAGCGCGGTATACCAACCCACATGGCCCACTACCT	3900
Db	3841	ATCCATCTTCTTCTCAGCTTCAGCGCGGTATACCAACCCACATGGCCCACTACCT	3900
Qy	3901	CAGGCTCATGTACAGTCAAGAAATGGTTCTCTCATCCAACTGCCCATGCGCCAATGATG	3960
Db	3901	CAGGCTCATGTACAGTCAAGAAATGGTTCTCTCATCCAACTGCCCATGCGCCAATGATG	3960
Qy	3961	CTAATGAGCAGCAGCCACCGCGGGTCCCGAGCGCCCTCGCTCAAAAGTGCACTACAG	4020
Db	3961	CTAATGAGCAGCAGCCACCGCGGGTCCCGAGCGCCCTCGCTCAAAAGTGCACTACAG	4020
Qy	4021	CCCAATCCAGTCTCGCAACAGCAGCTTTTCCCTTATATGAGCACCCTTCAGTACAAAGCC	4080
Db	4021	CCCAATCCAGTCTCGCAACAGCAGCTTTTCCCTTATATGAGCACCCTTCAGTACAAAGCC	4080
Qy	4081	CACCAACAGCAGTGTAGGTCCTTGGAGGAAACCGAAAGGCCAAATTTCCCTCTCTC	4140
Db	4081	CACCAACAGCAGTGTAGGTCCTTGGAGGAAACCGAAAGGCCAAATTTCCCTCTCTC	4140
Qy	4141	CCCTTCTACTGCTTCTACCACTGGAAGCAAGAAACTAGAAATTTCAATTTTGTGTTT	4200
Db	4141	CCCTTCTACTGCTTCTACCACTGGAAGCAAGAAACTAGAAATTTCAATTTTGTGTTT	4200
Qy	4201	TAAATATATATGTTGATTTCTTGTAACTCCAAATAGGAATGCTTAACAGTTTCACTTGCG	4260
Db	4201	TAAATATATATGTTGATTTCTTGTAACTCCAAATAGGAATGCTTAACAGTTTCACTTGCG	4260
Qy	4261	TGGAAGATACCTTGGACCGAGTAGAGGCAATTTAGGAACTTTGGGGCTATTCCCATTAATCCA	4320

```
Db 4261 TGAAGATACCTGGACCGAGTAGAGCAATTTAGGAACTTGGGGCTATTTCCTATAATTC 4320
QY 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
QY 4381 ATTTTAAATAACCCCTTGAAGTCAATGAAACATCATGCTAGCAAAAGAAAGTAAACAGAGT 4440
Db 4381 ATTTTAAATAACCCCTTGAAGTCAATGAAACATCATGCTAGCAAAAGAAAGTAAACAGAGT 4440
QY 4441 GATCTTCTGCTGCTTACTGCTAAAAAATAAAAAAATAAAAAA 4481
Db 4441 GATCTTCTGCTGCTTACTGCTAAAAAATAAAAAAATAAAAAA 4481

RESULT 6
HSU70323 4481 bp mRNA linear PRI 20-NOV-1996
LOCUS Human ataxin-2 (SCA2) mRNA, complete cds.
DEFINITION U70323
ACCESSION U70323
VERSION U70323.1 GI:1679683
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Pulst, S.-M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X.-N.,
Lopes-Cendes, I., Pearlman, S., Starkman, S., Orozco-biaz, G.,
Lunkes, A., DeJong, P., Rouleau, G.A., Auburger, G., Korenberg, J.R.,
Figueras, C. and Sahba, S.
TITLE Moderate expansion of a normally biallelic trinucleotide repeat in
spinocerebellar ataxia type 2
JOURNAL Nat. Genet. 14 (3), 269-276 (1996)
PUBMED 8896555
REFERENCE 2 (bases 1 to 4481)
AUTHORS Pulst, S.-M.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1996) Medicine, Cedars-Sinai, 8700 Beverly Blvd.,
Los Angeles, CA 90048, USA
FEATURES
source
1. 4481
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24.1"
1. 4481
/genes="SCA2"
/genes="SCA2"
/standard_name="spinocerebellar ataxia type 2"
/codon_start=1
/product="ataxin-2"
/protein_id="AAB19200.1"
/db_xref="GI:1679684"
/translation="MRSAAAPRSPAVATESRFRFAARWPFWSLQRPARRSGRGGG
AAGPVPSPAPPPGPPPPSRQSSPPSASDCPSGNGNGGAPRPSRLILGLGPPR
PFVVLLPLASGAPPAATRASPLGRASPPRSGVSLAKPAPCPRPAPCEPVYGLT
MSLKPQQQQQQQQQQQQQQQQQQPPAAANVRKPGSGLLASPAAPSPSSSV
SSSAPSVASVATSGRPPGLGRNKNKLPQSTISPDGIYANRMVHLITSVVG
SKCEVQVNGGIYGVGFKYSPKCDLVLDAAHEKSTSSSGPKREIMESILPKCDF
VVQPKDMSYAKRAFTDAISAKVNGRHKEDLPMDAGELTANEELEALENDVS
NGHPDNDPRYENGVWSTYDSSLSSTVPLERDSEBEFLKREARANQLAEEISS
AQYKRALVENDRSEBEKTAQRNRSEREGHSINTRENKIYPQQRNREVISWGG
QNSPRMPQPGSMPSRSTSHSDFNPNSSDQRVVNGVWPSPCPSPSPSPSPRY
QSGPNSLPQPAATPTPPSPRPPRPPSHSAGSPAPVSTMPKRMSEGPSPRSP
KAQHPNHRVSRAGRSISGLLEFVSHNPPEAATPPVARTSPSGGTWSSVSVSPRL
SKTHPRSPRONSGNTPSGPVLASPAQGIIPTEAVAMPDPAASPTPASPNRAVT
PSSEAKDSRLQDRQNSPAGNENIKPNETSPSFKENKGISPVVSEHRKQIDDLKK
FKNDFRLQPSSTSEMDQLLNKREGEKSKDLIDKIEPSAKDSFIENSNSNCTSGSS
```

```
KNPSPISTSLNTEHKRGPEVTSQVOTSSPACKOBKDBKEEKDAAEVOVRKSTLN
PNAKEPNRPSFQPKPSTPTSPRPAQPSBVMVCHQOPTPVYTPVCFANMXYVPP
VSPGVQPLPIMPMPVNPVQAKTYTRAVPNMFQORODQHQSMMHPASAGPPILAATP
PAYSTQVAYVSPQPNVQVHVPYOSQHPHYSPVIOGNARMMAPPTTHAOPGLVS
SSATQYGAHEQTHAMACPKLPYKNETSPFYFAISTGLSQAQYAHNATLHPHTPH
QPSATPTGQOQSGHGHSPAPSVQHQHQAALHLASPOQSAIYHAGLAPTPPSM
TPASNTQSPNSFPAAQOTVFTIHPSHVQPAYTNPHPMAHVPOAHVQSMVPSHPTAH
APMMLMTQPPGGPQAALQALQPIPVSTTAHFPYMYTHPSVQAHHQOOL"
ORIGIN
Query Match 100.0%; Score 4481; DB 8; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCTCCCGCCCCCGG 60
Db 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCTCCCGCCCCCGG 60
QY 61 GCCACCTCAGCTTCTGCTTCCGTCGTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
Db 61 GCCACCTCAGCTTCTGCTTCCGTCGTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
QY 121 CACCTCCGCTCCCAACCCCGCGCTCCGCGCGCGCGCTCCGATCGGTCAGCGCGCGCA 180
Db 121 CACCTCCGCTCCCAACCCCGCGCTCCGCGCGCGCGCTCCGATCGGTCAGCGCGCGCA 180
QY 181 GCTCTTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCCGCGAGCCAGTGGCCCC 240
Db 181 GCTCTTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCCGCGAGCCAGTGGCCCC 240
QY 241 GGGTGGCGCTCCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 GGGTGGCGCTCCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CCGGACCGTATCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 CCGGACCGTATCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CAGAGTCCGCTTCCCTCCGCTCAGACTGTTTGTAGCAACCGCGCGCGCGCGCG 420
Db 361 CAGAGTCCGCTTCCCTCCGCTCAGACTGTTTGTAGCAACCGCGCGCGCGCGCG 420
QY 421 TTTTCGCGCGCGCTCCCGCGCGCTTCTTGTCTCGCGCGCGCGCGCGCGCGCG 480
Db 421 TTTTCGCGCGCGCTCCCGCGCGCTTCTTGTCTCGCGCGCGCGCGCGCGCGCG 480
QY 481 GTCTTCTCTCCCTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 481 GTCTTCTCTCCCTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 CTTCGCGCGCGCTCCCGCGCGCTTCCGCGCGCTTCTTGTGCGCGCGCGCGCGCG 600
Db 541 CTTCGCGCGCGCTCCCGCGCGCTTCCGCGCGCTTCTTGTGCGCGCGCGCGCGCG 600
QY 601 TGTTCGCGCGCGCGCTGCGAGCGGTGATGGCGCGCTTCCACCATGTCGCTGAAGCCCCAG 660
Db 601 TGTTCGCGCGCGCGCTGCGAGCGGTGATGGCGCGCTTCCACCATGTCGCTGAAGCCCCAG 660
QY 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
Db 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
QY 721 CAGCCGCGCGCGCGCTGCTCCTTCTTGTGCGCGCGCGCGCGCGCGCGCGCTTCTAGCGTCG 780
Db 721 CAGCCGCGCGCGCGCTGCTCCTTCTTGTGCGCGCGCGCGCGCGCGCGCGCTTCTAGCGTCG 780
QY 781 CCGCGCGCGCGCGCTTCCGCGCTTCTTGTGCGCGCGCGCGCGCGCGCGCGCTTCTAGCGTCG 840
Db 781 CCGCGCGCGCGCGCTTCCGCGCTTCTTGTGCGCGCGCGCGCGCGCGCGCGCTTCTAGCGTCG 840
QY 841 TCCTCGGTGTCGCGCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 TCCTCGGTGTCGCGCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
```

QY	901	AGTAA	CAAAAGGACTGCCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATATGAGG	960
DB	901	AGTAA	CAAAAGGACTGCCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATATGAGG	960
QY	961	ATGGTT	CCATATACTTACATCAGTTCTTGGCTCCAAATGCTGAAGTACCAAGTGAAAAATGGA	1020
DB	961	ATGGTT	CCATATACTTACATCAGTTCTTGGCTCCAAATGCTGAAGTACCAAGTGAAAAATGGA	1020
QY	1021	GGTATAT	GAAGGAGTTTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTTGATGCC	1080
DB	1021	GGTATAT	GAAGGAGTTTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTTGATGCC	1080
QY	1081	GCACAT	GAGAAAAGTACAGAAATCCAGTTTCGGGGCGGAAACGTGAAGAAATAATCGAGAGT	1140
DB	1081	GCACAT	GAGAAAAGTACAGAAATCCAGTTTCGGGGCGGAAACGTGAAGAAATAATCGAGAGT	1140
QY	1141	ATTTTGT	TCAAATGTTCCAGACTTTTGTGTGCTACAGTTTAAAGATATGGACTCCAGTTAT	1200
DB	1141	ATTTTGT	TCAAATGTTCCAGACTTTTGTGTGCTACAGTTTAAAGATATGGACTCCAGTTAT	1200
QY	1201	GCAAAA	GAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAACACAAA	1260
DB	1201	GCAAAA	GAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAACACAAA	1260
QY	1261	GAGAAG	CACTGGAGCCCTGGATGCAAGTGAACTCACAGCCAATGAGGAACTTTGAGGCT	1320
DB	1261	GAGAAG	CACTGGAGCCCTGGATGCAAGTGAACTCACAGCCAATGAGGAACTTTGAGGCT	1320
QY	1321	TTGGAAA	TATGACGTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
DB	1321	TTGGAAA	TATGACGTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
QY	1381	AAATTAT	GGTGTAGTGTCTACCTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA	1440
DB	1381	AAATTAT	GGTGTAGTGTCTACCTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA	1440
QY	1441	AGAGATA	AACTCAGAAGAATTTTTTAAACCGGGAACCAAGGGCAAAACCCAGTTAGCAGAA	1500
DB	1441	AGAGATA	AACTCAGAAGAATTTTTTAAACCGGGAACCAAGGGCAAAACCCAGTTAGCAGAA	1500
QY	1501	ATTGAGT	CAAGTGCACGATCAAAAGCTCGAGTGGCCCTGGGAAATGATATGAGAGTGAG	1560
DB	1501	ATTGAGT	CAAGTGCACGATCAAAAGCTCGAGTGGCCCTGGGAAATGATATGAGAGTGAG	1560
QY	1561	GAAGAAA	AAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCAACAGCATAAAC	1620
DB	1561	GAAGAAA	AAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCAACAGCATAAAC	1620
QY	1621	ACTAGG	GAAAAATAATATATTTCTCTGGACAAAGAAATAGAGAAGTCATATCTCGGGGA	1680
DB	1621	ACTAGG	GAAAAATAATATATTTCTCTGGACAAAGAAATAGAGAAGTCATATCTCGGGGA	1680
QY	1681	AGTGGG	GACAGAAATTCACCGCGTATGGGCCAGCTGATCGGGCTCATATGCCATCAAGA	1740
DB	1681	AGTGGG	GACAGAAATTCACCGCGTATGGGCCAGCTGATCGGGCTCATATGCCATCAAGA	1740
QY	1741	TCCACT	CTTCACACTTCAGATTTTCAACCCGAAATTCGTGTTTCAGACCAAGAGTAGTTAAT	1800
DB	1741	TCCACT	CTTCACACTTCAGATTTTCAACCCGAAATTCGTGTTTCAGACCAAGAGTAGTTAAT	1800
QY	1801	GGAGGT	GTTCCTGGCCATCGCCTTGCCCATCTCTCTCTCGGCCCAACCTTCTCGCTAC	1860
DB	1801	GGAGGT	GTTCCTGGCCATCGCCTTGCCCATCTCTCTCTCGGCCCAACCTTCTCGCTAC	1860
QY	1861	CAGTCAG	TCCCAACTCTTTCACCTCGGGCAGCAACCCCTACACGGCGCCCTCCAGG	1920
DB	1861	CAGTCAG	TCCCAACTCTTTCACCTCGGGCAGCAACCCCTACACGGCGCCCTCCAGG	1920
QY	1921	CCCCCT	TCGGGCCATTCAGACCCCGCTCTCACCCCTCTGCTCATGTTCTTCAGCTCCT	1980
DB	1921	CCCCCT	TCGGGCCATTCAGACCCCGCTCTCACCCCTCTGCTCATGTTCTTCAGCTCCT	1980

Qy	1981	GTCTCTACTATGCTCTAAACGCATGTCTTTACAGAAAGGGCCTCTCAAGAGATGTCTCCCAAAAGGCC	2040
Db	1981	GTCTCTACTATGCTCTAAACGCATGTCTTTACAGAAAGGGCCTCTCAAGAGATGTCTCCCAAAAGGCC	2040
Qy	2041	CAGCGACATCCTCGAAATCACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCGCTA	2100
Db	2041	CAGCGACATCCTCGAAATCACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCGCTA	2100
Qy	2101	GAATTTTGATTCGCCAACCCACCCAGTGAAGCAGCTACTCCTCCAGTAGCAAGGACCCAGT	2160
Db	2101	GAATTTTGATTCGCCAACCCACCCAGTGAAGCAGCTACTCCTCCAGTAGCAAGGACCCAGT	2160
Qy	2161	CCCTCGGGGGGAAAGCTGGTTCATCAGTGGTTCAGTGGGGTTCCAAAGATTTATCCCTTAAAACT	2220
Db	2161	CCCTCGGGGGGAAAGCTGGTTCATCAGTGGTTCAGTGGGGTTCCAAAGATTTATCCCTTAAAACT	2220
Qy	2221	CATAGACCCAGGTCTCCACACAGACAGATATTGGAAATACCCCCAGTGGGCCAGTTCTTT	2280
Db	2221	CATAGACCCAGGTCTCCACACAGACAGATATTGGAAATACCCCCAGTGGGCCAGTTCTTT	2280
Qy	2281	GCCTTCTCCCAAGCTGGTATTATTTCCAACTGAAAGCTGTTGCCATGCCCTATTTCAGAGTGC	2340
Db	2281	GCCTTCTCCCAAGCTGGTATTATTTCCAACTGAAAGCTGTTGCCATGCCCTATTTCAGAGTGC	2340
Qy	2341	TCTCTTAAGCTGTCTGTCTGTCATCGAAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA	2400
Db	2341	TCTCTTAAGCTGTCTGTCTGTCATCGAAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA	2400
Qy	2401	GATTTCAGGCTTCAAGATCAGAGCGAGACTCTCCTGCAGGGAATAAGAAATATTTAAA	2460
Db	2401	GATTTCAGGCTTCAAGATCAGAGCGAGACTCTCCTGCAGGGAATAAGAAATATTTAAA	2460
Qy	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGMAAAACAAAGATATCACACAGTTGTT	2520
Db	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGMAAAACAAAGATATCACACAGTTGTT	2520
Qy	2521	TCTGAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTAAGAAATGATTTTAGGTTACAG	2580
Db	2521	TCTGAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTAAGAAATGATTTTAGGTTACAG	2580
Qy	2581	CCAAGTTCTACTTCTGAACTCTATGGATCAACTACTAAACAAAAATAGAGAGGAGAAAAA	2640
Db	2581	CCAAGTTCTACTTCTGAACTCTATGGATCAACTACTAAACAAAAATAGAGAGGAGAAAAA	2640
Qy	2641	TCAAGAGATTGATCAAAAGACAAAATTTGAACCAAGTCTAAGGATTCCTTTCATTGAAAT	2700
Db	2641	TCAAGAGATTGATCAAAAGACAAAATTTGAACCAAGTCTAAGGATTCCTTTCATTGAAAT	2700
Qy	2701	AGCAGCAGCAACTGTATCCAGTGGCAGCAGCAAGCCGAATAGCCCCCAGCATTTCCCCTTCA	2760
Db	2701	AGCAGCAGCAACTGTATCCAGTGGCAGCAGCAAGCCGAATAGCCCCCAGCATTTCCCCTTCA	2760
Qy	2761	ATACTTAGTAAACGAGGACACAGAGGGGACCTGAGTCACTTTCCCAAGGGGTTTCAGACT	2820
Db	2761	ATACTTAGTAAACGAGGACACAGAGGGGACCTGAGTCACTTTCCCAAGGGGTTTCAGACT	2820
Qy	2821	TCCAGCCCGCATGTATAACAGAGAAAGACGATAAGGAAGAGAAAGACGAGCTGAG	2880
Db	2821	TCCAGCCCGCATGTATAACAGAGAAAGACGATAAGGAAGAGAAAGACGAGCTGAG	2880
Qy	2881	CAAGTTAGGAAATCAACATTTGAAATCCCAATGCAAAAGAGTTTCAACCCAGCTTCTCTCT	2940
Db	2881	CAAGTTAGGAAATCAACATTTGAAATCCCAATGCAAAAGAGTTTCAACCCAGCTTCTCTCT	2940
Qy	2941	CAGCCAAAGCCTTCTACTACCCCAACTTTCACCTCGGCTCAAGGACAACTAGGCCCATCT	3000
Db	2941	CAGCCAAAGCCTTCTACTACCCCAACTTTCACCTCGGCTCAAGGACAACTAGGCCCATCT	3000
Qy	3001	ATGGTGGTTCATCAACAGCCAACTCCAGTTTATCTCAGGCTGTTGTTTGCACCAAAAT	3060
Db	3001	ATGGTGGTTCATCAACAGCCAACTCCAGTTTATCTCAGGCTGTTGTTTGCACCAAAAT	3060
Qy	3061	ATGATGTATCCAGTCCAGTGGCCAGCCAGGGGTCACACTTTTATACCCAAATACCTATGACG	3120

```
Db 3061 ATGATGTATCCAGTCCAGTGAGCCAGCGGTGCAACCTTTATACCAATACCTATGAGC 3120
Qy 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCAAACAGCGG 3180
Db 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCAAACAGCGG 3180
Qy 3181 CAAGACCAGCATCATCAGAGTGCATGATGACACCCAGCGTCAAGCAGCGGGCCACCGATT 3240
Db 3181 CAAGACCAGCATCATCAGAGTGCATGATGACACCCAGCGTCAAGCAGCGGGCCACCGATT 3240
Qy 3241 GCAGCCACCCACCCAGCTTACTCCAGCAATATGTTGCCCTACAGTCCCTCAGCAGTTCCTCA 3300
Db 3241 GCAGCCACCCACCCAGCTTACTCCAGCAATATGTTGCCCTACAGTCCCTCAGCAGTTCCTCA 3300
Qy 3301 AATCAGCCCTTTGTTTCAGCATGTGCAATATATCATGCTCTCAGCATCTCATGCTATAGT 3360
Db 3301 AATCAGCCCTTTGTTTCAGCATGTGCAATATATCATGCTCTCAGCATCTCATGCTATAGT 3360
Qy 3361 CCTGTAATACAGGGTAACTGTAAGATGATGGCAACCAACACAGCCCGAGCTGGTTTA 3420
Db 3361 CCTGTAATACAGGGTAACTGTAAGATGATGGCAACCAACACAGCCCGAGCTGGTTTA 3420
Qy 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCATGCGATGTATGATGT 3480
Db 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCATGCGATGTATGATGT 3480
Qy 3481 CCCAAATTAACCATACAAAGGAGCAAGCCCTTTCTTCTACTTTGGCCATTTCCACGGGC 3540
Db 3481 CCCAAATTAACCATACAAAGGAGCAAGCCCTTTCTTCTACTTTGGCCATTTCCACGGGC 3540
Qy 3541 TCCCTTGCTCAGCAGTATGGCAACCTTAACGCTACCCCTGCAACCCACATACCTCAGACCT 3600
Db 3541 TCCCTTGCTCAGCAGTATGGCAACCTTAACGCTACCCCTGCAACCCACATACCTCAGACCT 3600
Qy 3601 CAGCTTTCAGTACCCCTGACAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3660
Db 3601 CAGCTTTCAGTACCCCTGACAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3660
Qy 3661 CCAGTCTCTGTTTCCAGCACCATCAGCAGCAGCGCCCGCCAGGCTCTCCATCTGGCCAGTCA 3720
Db 3661 CCAGTCTCTGTTTCCAGCACCATCAGCAGCAGCGCCCGCCAGGCTCTCCATCTGGCCAGTCA 3720
Qy 3721 CAGCAGCAGTACGCAATTTACAGCGGGGCTTGGCCCAACTCCACCCCTCCATGACACT 3780
Db 3721 CAGCAGCAGTACGCAATTTACAGCGGGGCTTGGCCCAACTCCACCCCTCCATGACACT 3780
Qy 3781 GCCTCCAAACAGCAGTCCGACAGATAGTTTCCAGCAGCAACACAGACTGCTTTTACG 3840
Db 3781 GCCTCCAAACAGCAGTCCGACAGATAGTTTCCAGCAGCAACACAGACTGCTTTTACG 3840
Qy 3841 ATCCATCTCTTCTCAGCGTTTCCAGCGGGGTATACCAACCCCAATGCGCCCACTGACCT 3900
Db 3841 ATCCATCTCTTCTCAGCGTTTCCAGCGGGGTATACCAACCCCAATGCGCCCACTGACCT 3900
Qy 3901 CAGGCTCATGTA CAGTCAGGAATGGTTCCTTCTCATCCAACTGCCCATGCGCCCAATGATG 3960
Db 3901 CAGGCTCATGTA CAGTCAGGAATGGTTCCTTCTCATCCAACTGCCCATGCGCCCAATGATG 3960
Qy 3961 CTAAATGACGACAGCAGCAGCCCGGGTCCAGCGCGCCCTCGCTCAAAAGTGCATACAG 4020
Db 3961 CTAAATGACGACAGCAGCAGCCCGGGTCCAGCGCGCCCTCGCTCAAAAGTGCATACAG 4020
Qy 4021 CCCATTTCCAGTCTCGAACAACAGCGCATTTTCCCTTATATGAGCAGCCCTTCAGTACAGCC 4080
Db 4021 CCCATTTCCAGTCTCGAACAACAGCGCATTTTCCCTTATATGAGCAGCCCTTCAGTACAGCC 4080
Qy 4081 CACCAACAAACAGCAGTGTGAAGGCTGCTGGAGGAAACCGAAAGGCGCAATTCCTTCCTC 4140
Db 4081 CACCAACAAACAGCAGTGTGAAGGCTGCTGGAGGAAACCGAAAGGCGCAATTCCTTCCTC 4140
Qy 4141 CCTTCTACTGCTTCTACCAACTGGAGCAGCAAGAAACTAGAAATTCATTTATTTGTTTT 4200
```

```
Db 4141 CCTTCTACTGCTTCTACCAACTGGAAGCACAGAAAACTAGAATTTCAATTTATTTGTTTT 4200
Qy 4201 TAAATATATATGTTGATTTCTTGTAAACATCCATAGGAATGCCTAACAGTTCACTTGCAG 4260
Db 4201 TAAATATATATGTTGATTTCTTGTAAACATCCATAGGAATGCCTAACAGTTCACTTGCAG 4260
Qy 4261 TGGAAAGATACTTGGACCCAGTAGAGGCATTTAGGAACCTTGGGGGCTATTTCCTATAATTCCA 4320
Db 4261 TGGAAAGATACTTGGACCCAGTAGAGGCATTTAGGAACCTTGGGGGCTATTTCCTATAATTCCA 4320
Qy 4321 TATGCTGTTT CAGAGTCCCGCAGTACCCCGAGCTCTGCTTGC CGAACTTGAAGTATTATTT 4380
Db 4321 TATGCTGTTT CAGAGTCCCGCAGTACCCCGAGCTCTGCTTGC CGAACTTGAAGTATTATTT 4380
Qy 4381 ATTTTAAATACCCCTTGAAGTCTATGAACATCAGCTAGCAAAAGAGTAACAAGACT 4440
Db 4381 ATTTTAAATACCCCTTGAAGTCTATGAACATCAGCTAGCAAAAGAGTAACAAGACT 4440
Qy 4441 GATTTCTGCTGCTATTACTGCTTAAAAAATAAAAAA 4481
Db 4441 GATTTCTGCTGCTATTACTGCTTAAAAAATAAAAAA 4481
```

## RESULT 7

```
C0727533 4479 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 13467 from Patent WO02068579.
DEFINITION C0727533
ACCESSION C0727533
VERSION C0727533.1 GI:42293778
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 13467 06-SEP-2002;
PE Corporation (NV) (US)
FEATURES
source 1..4479
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

## ORIGIN

```
Query Match 99.3%; Score 4450.6; DB 6; Length 4479;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4475; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
Qy 1 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCTCTACAGTGTCCCTCCCGCCCCGGG 60
Db 1 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCTCTACAGTGTCCCTCCCGCCCCGGG 60
Qy 61 GCCACCTCAGTTCCTGCTTCCGTTCTGACCCCTCCGACTTCCGTTAAAGAGTCCCTATCCG 120
Db 61 GCCACCTCAGTTCCTGCTTCCGTTCTGACCCCTCCGACTTCCGTTAAAGAGTCCCTATCCG 120
Qy 121 CACCTCCGTCCTCAGCCCGCGCTCTGGCGCGCCGCTCCGATGCGCTCAGCGCGCGCA 180
Db 121 CACCTCCGTCCTCAGCCCGCGCTCTGGCGCGCCGCTCCGATGCGCTCAGCGCGCGCA 180
Qy 181 GCTCTCGAGTCCCGCGGTGCGCACCGAGTCTTCGCGCTTCGCCCGAGCAGGTGGCCC 240
Db 181 GCTCTCGAGTCCCGCGGTGCGCACCGAGTCTTCGCGCTTCGCCCGAGCAGGTGGCCC 239
Qy 241 GGGTGGCGCTCGTCCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGTGGCGCGCC 300
Db 240 GGGTGGCGCTCGTCCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGTGGCGCGCC 299
```

QY 301 CCGGACCGTATCCCTCCGCGCCCTCCCGCCGCGCCGCGCCCTCCCTCCCGG 360  
Db |||||  
QY 300 CCGGACCGTATCCCTCCGCGCCCTCCCGCCGCGCCGCGCCCTCCCTCCCGG 359  
Db |||||  
QY 361 CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCG 420  
Db |||||  
QY 360 CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCG 419  
QY 421 TTTCCGCGCGGCTCCGCGGCGCTCTTGGTCTCGCGGCGCTCCCGCCCTTCGTCGTC 480  
Db |||||  
QY 420 TTTCCGCGCGGCTCCGCGGCGCTCTTGGTCTCGCGGCGCTCCCGCCCTTCGTCGTC 479  
QY 481 GTCTTCTCCCTCCGCGCGCGCGCGCCCTCCGCGCGCGCGCAACCGCGCGCTCCCGG 540  
Db |||||  
QY 480 CTCCTTCTCCCTCCGCGCGCGCGCGCCCTCCGCGCGCGCGCAACCGCGCGCTCCCGG 539  
QY 541 CTCGCGCGCGGCTCCGCGCGCTCCGCGGCTCTCTTGGCGCGCGCGGCTCCCGGC 600  
Db |||||  
QY 540 CTCGCGCGCGGCTCCGCGCGCTCCGCGGCTCTCTTGGCGCGCGCGGCTCCCGGC 599  
QY 601 TGTCCCGCGCGGCTCGAGCGGTGATGGGCCCCCTCAACATGCTCGCTGAAGCCCCAG 660  
Db |||||  
QY 600 TGTCCCGCGCGGCTCGAGCGGTGATGGGCCCCCTCAACATGCTCGCTGAAGCCCCAG 659  
QY 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720  
Db |||||  
QY 660 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 719  
QY 721 CAGCCCGCGCGCGGCTGCCAATGTCGCAAGCCCGCGCGCGGCTTCTAGCGTCG 780  
Db |||||  
QY 720 CAGCCCGCGCGCGGCTGCCAATGTCGCAAGCCCGCGCGCGGCTTCTAGCGTCG 779  
QY 781 CCCGCGCGCGGCTTTCGCGGCTCTCGTCTCGGTCCTCGTCTCGGCAACGCGCTCCC 840  
Db |||||  
QY 780 CCCGCGCGCGG - CTTTCGCGGCTCTCGTCTCGGTCCTCGTCTCGGCAACGCGCTCCC 838  
QY 841 TCCTCGGTCGCGCGGACCTCCGCGCGCGGAGCGCGGCTTTCGCGAGGTCGAAC 900  
Db |||||  
QY 839 TCCTCGGTCGCGCGGACCTCCGCGCGCGGAGCGCGGCTTTCGCGAGGTCGAAC 898  
QY 901 AGTAAACAAAGGACTCGCTCAGCTACGATTTCTTTTGATGGAATCTAGCAAAATAGG 960  
Db |||||  
QY 899 AGTAAACAAAGGACTCGCTCAGCTACGATTTCTTTTGATGGAATCTAGCAAAATAGG 958  
QY 961 ATGGTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAAGGA 1020  
Db |||||  
QY 959 ATGGTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAAGGA 1018  
QY 1021 GGTATATATGAGGAGTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080  
Db |||||  
QY 1019 GGTATATATGAGGAGTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1078  
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGCGAAACGTGAAGAAATATGAGAGT 1140  
Db |||||  
QY 1079 GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGCGAAACGTGAAGAAATATGAGAGT 1138  
QY 1141 ATTTTGTTCAAATGTTACAGCTTTGTTGGTACAGTTTAAAGATPATGGACTCCAGTTAT 1200  
Db |||||  
QY 1139 ATTTTGTTCAAATGTTACAGCTTTGTTGGTACAGTTTAAAGATPATGGACTCCAGTTAT 1198  
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGAAGTGAATGGCGCAACAAA 1260  
Db |||||  
QY 1199 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTGAAGTGAATGGCGCAACAAA 1258  
QY 1261 GAGAAAGACCTGGAGCCCTGGGATCAGGTGAATCAGAGCAATGAGGAACTTGAGGCT 1320  
Db |||||  
QY 1259 GAGAAAGACCTGGAGCCCTGGATCAGGTGAATCAGAGCAATGAGGAACTTGAGGCT 1318  
QY 1321 TTGAAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATAATGAAGAA 1380  
Db |||||  
QY 1319 TTGAAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATAATGAAGAA 1378  
QY 1381 AATTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA 1440

Db 1379 AATTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA 1438  
QY 1441 AGAGATAACTCAGAAGAAATTTTAAAAACGGGAACAAAGGCAAAACAGTTAGCAAGAA 1500  
Db 1439 AGAGATAACTCAGAAGAAATTTTAAAAACGGGAACAAAGGCAAAACAGTTAGCAAGAA 1498  
QY 1501 ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATAGGAGTGAG 1560  
Db 1499 ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATAGGAGTGAG 1558  
QY 1561 GAAGAAAAATACAGCAGAGTTTCAGAGAAATTCAGAGTGAACGTGAGGGGCACAGATAAAC 1620  
Db 1559 GAAGAAAAATACAGCAGAGTTTCAGAGAAATTCAGAGTGAACGTGAGGGGCACAGATAAAC 1618  
QY 1621 ACTAGGGAAAAATAATATATTTCTCTCTGGAACAAAGAAATAGAGAGTTCATATCTCTGGGA 1680  
Db 1619 ACTAGGGAAAAATAATATATTTCTCTCTGGAACAAAGAAATAGAGAGTTCATATCTCTGGGA 1678  
QY 1681 AGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCTTGGATCGGCTCCATGCCATCAAGA 1740  
Db 1679 AGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCTTGGATCGGCTCCATGCCATCAAGA 1738  
QY 1741 TCCACTTCTCACACTTCAGATTTTCAACCCGAATTCAGTGGTTTCAGACCAAGAGTAGTTAAT 1800  
Db 1739 TCCACTTCTCACACTTCAGATTTTCAACCCGAATTCAGTGGTTTCAGACCAAGAGTAGTTAAT 1798  
QY 1801 GGAGTGTTCCTCGCCATCGCCTTGGCCATCTCTCTCTCGCCACCTTCTCGCTAC 1860  
Db 1799 GGAGTGTTCCTCGCCATCGCCTTGGCCATCTCTCTCTCGCCACCTTCTCGCTAC 1858  
QY 1861 CAGTCAGTCCCAACTCTCTTTCACCTCGGCGAGCCACCTTACAGCGCGCCCTCCAGG 1920  
Db 1859 CAGTCAGTCCCAACTCTCTTTCACCTCGGCGAGCCACCTTACAGCGCGCCCTCCAGG 1918  
QY 1921 CCCCTCGCGGCGCATCCAGACCCGCTCTCACCCCTCTGCTCATGTTCTTCAGTCTCT 1980  
Db 1919 CCCCTCGCGGCGCATCCAGACCCGCTCTCACCCCTCTGCTCATGTTCTTCAGTCTCT 1978  
QY 1981 GTCTCTACTATGCTTAAACGATGCTTTTCAAGAGGCGCTTCAAGGATGTCTCCAAAGGCC 2040  
Db 1979 GTCTCTACTATGCTTAAACGATGCTTTTCAAGAGGCGCTTCAAGGATGTCTCCAAAGGCC 2038  
QY 2041 CAGCGACATCTCGAAATCAGAGGTTTCTGCTGGAGGGGTTTCCATATTCAGTGGGCTA 2100  
Db 2039 CAGCGACATCTCGAAATCAGAGGTTTCTGCTGGAGGGGTTTCCATATTCAGTGGGCTA 2098  
QY 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAGGACCACT 2160  
Db 2099 GAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAGGACCACT 2158  
QY 2161 CCCTCGGGGGGAAAGTGTCTCATGTTGAGTGGGTTTCAAGATTTATCCCTTAAACT 2220  
Db 2159 CCCTCGGGGGGAAAGTGTCTCATGTTGAGTGGGTTTCAAGATTTATCCCTTAAACT 2218  
QY 2221 CATAGACCCAGGTCTCCAGACAGAAAGATTTGAAATACCCCGAGTGGGCGGCTTCTT 2280  
Db 2219 CATAGACCCAGGTCTCCAGACAGAAAGATTTGAAATACCCCGAGTGGGCGGCTTCTT 2278  
QY 2281 GCTTCTCCCAAGCTGGTATTTTCCAACTGAAGCTGTTGCCATGCCCTATTTCCAGCTGCA 2340  
Db 2279 GCTTCTCCCAAGCTGGTATTTTCCAACTGAAGCTGTTGCCATGCCCTATTTCCAGCTGCA 2338  
QY 2341 TCTCTACGCTGTAGTCTGCTGCAACAGAGCTGTTACCCCTTCTAGTGAGGCTAAA 2400  
Db 2339 TCTCTACGCTGTAGTCTGCTGCAACAGAGCTGTTACCCCTTCTAGTGAGGCTAAA 2398  
QY 2401 GATTCCAGGCTTCAAGATCAGAGCAGAACTCTCTCGAGGGAATAAGAAAAATTTAAA 2460  
Db 2399 GATTCCAGGCTTCAAGATCAGAGCAGAACTCTCTCGAGGGAATAAGAAAAATTTAAA 2458  
QY 2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGTAAGCAAGAGGTATATACCGAGTCTT 2520















Db	3240	CTCTCCATCTGGCCAGTCCACAGCAGCAGT	CAGCCATTTACACGCGGGGCTTGCGCAA	3299
QY	3761	CTCCACCTTCATGACACTGCTCTCAACACG	CAGTGCACAGATAATAGTTTCCAGCAG	3820
Db	3300	CTCCACCTTCATGACACTGCTCTCAACACG	CAGTGCACAGATAATAGTTTCCAGCAG	3359
QY	3821	CACACAGACTGCTCTTACGATCCATCTTCT	CACAGTTCACGCGGCGTATACCAACCCAC	3880
Db	3360	CACACAGACTGCTCTTACGATCCATCTTCT	CACAGTTCACGCGGCGTATACCAACCCAC	3419
QY	3881	CCACATGCCCCACACTACTCAGGCTCATGT	ACAGTACAGGAATGTTCTCTCATCAAA	3940
Db	3420	CCACATGCCCCACACTACTCAGGCTCATGT	ACAGTACAGGAATGTTCTCTCATCAAA	3479
QY	3941	CTGCCCATGCGCCAATGATGCTAATGACG	ACACGCCACCGCGGTCCTCCAGGCGGCC	4000
Db	3480	CTGCCCATGCGCCAATGATGCTAATGACG	ACACGCCACCGCGGTCCTCCAGGCGGCC	3539
QY	4001	TGCTCAAAAGTGCACTACAGCCCATTCAG	TCTCGACACAGCGCATTTCCCTATATGA	4060
Db	3540	TGCTCAAAAGTGCACTACAGCCCATTCAG	TCTCGACACAGCGCATTTCCCTATATGA	3599
QY	4061	CGACCTTCAGTACAAGCCACACCAACAG	CAGCTTGTAAAGCTGCGCTGGAGAACG	4120
Db	3600	CGACCTTCAGTACAAGCCACACCAACAG	CAGCTTGTAAAGCTGCGCTGGAGAACG	3659
QY	4121	AAAGGCCAAATTCCTCTCTCTTACTGCTT	CTACCAACTGGAAGCAGACAGAAACTAG	4180
Db	3660	AAAGGCCAAATTCCTCTCTCTTACTGCTT	CTACCAACTGGAAGCAGACAGAAACTAG	3719
QY	4181	AATTTCAATTAATTTGTTTTTAAATATAT	ATATGTTGTAATCTTGTAAACATCAATAG	4240
Db	3720	AATTTCAATTAATTTGTTTTTAAATATAT	ATATGTTGTAATCTTGTAAACATCAATAG	3779
QY	4241	TGCTAACAGTTTCACATTGAGTGGAGATA	CTTGGACCGAGTAGAGGCATTTAGGAATTG	4300
Db	3780	TGCTAACAGTTTCACATTGAGTGGAGATA	CTTGGACCGAGTAGAGGCATTTAGGAATTG	3839
QY	4301	GGGGCTATTCCATAATTCATATGCTGTTT	CAGAGTCCGCGAGGTACCCAGCTCTGTT	4360
Db	3840	GGGGCTATTCCATAATTCATATGCTGTTT	CAGAGTCCGCGAGGTACCCAGCTCTGTT	3899
QY	4361	GCCGAACTGGAGTATTATTTTAAATACCT	TGAAAGTCAATGACATCAGCTA	4420
Db	3900	GCCGAACTGGAGTATTATTTTAAATACCT	TGAAAGTCAATGACATCAGCTA	3959
QY	4421	GCAAAAGAGTAACAAGAGTATCTTCTGCT	ATTACTTGCTTAAACCAAAAAAAAAA	4480
Db	3960	GCAAAAGAGTAACAAGAGTATCTTCTGCT	ATTACTTGCTTAAACCAAAAAAAAAA	4019
QY	4481	A 4481		
Db	4020	A 4020		
RESULT 10				
AF041472				
LOCUS				
DEFINITION Mus musculus ataxin-2 (SCA2) mRNA, linear ROD 28-NOV-2001				
ACCESSION AF041472				
VERSION AF041472.1 GI:3005019				
KEYWORDS				
SOURCE Mus musculus (house mouse)				
ORGANISM				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;				
Sciurognathi; Murioidea; Muridae; Murinae; Mus.				
REFERENCE 1 (bases 1 to 4225)				
AUTHORS Nechiporuk, T., Huynh, D.P., Figueroa, K., Sahba, S., Nechiporuk, A. and Pulst, S.M.				
TITLE The mouse SCA2 gene: cDNA sequence, alternative splicing and protein expression				
JOURNAL Hum. Mol. Genet. 7 (8), 1301-1309 (1998)				

9668173	PUBMED	2 (bases 1 to 4225)	Nechiporuk, T., Figueroa, K., Sahba, S., Nechiporuk, A.V. and Pulst, S.M.
	REFERENCE		
	AUTHORS		
	TITLE	Direct Submission	
	JOURNAL	Submitted (07-JAN-1998) Medicine/Neurology, Cedars-Sinai Medical Center, 8700 Beverly Blvd., Los Angeles, CA 90048, USA	
	FEATURES	Location/Qualifiers	
	source	1. 4225	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10090"	
		/chromosome="12"	
		/map="12q23.1"	
	gene	1. 4225	
		/gene="SCA2"	
	CDS	27..3884	
		/gene="SCA2"	
		/codon_start=1	
		/product="ataxin-2"	
		/protein_id="AAC09275.1"	
		/db_xref="GI:3005020"	
		/translation="MRSSTAQVORPAAGDEPRPAGMAARRSLPRTARGRGCAVA YPSGPPRPGAPPRGPRSPPCASDCFGNGHGASRGRRLLCVCGPRPPVVVLL ALAPATPARACGLPGVRSAPRSGVSSARPACGPRPACPEVYGLTMSLUKPOPPP APATKRPKGLLSPGASAPASAAVTSASVVPAPAPVASSAAAGGRLGRLGRNS SKGLPQPTISFDGIYANVRVHILTSVSGKEVQKNGGIYEGVFYKTSKCLDLVD AAEKEDJLEPMDAGELTASELELNDVNGMDPNDMPRYNEENVGVSTVDSSLSSTV VPLERDNEBFLKREARANQLEBIESSAQYKARVALENDRESEEEKYATVQRNCSR EHGPNTRDNKIIPPGORNREVLSWGSQSRSGRMQPGFGSPMSRAASHTSDPNNA GSDRVNGVGVPMFPCPSHSSRPPRYQSPNSLPRAATHTRPPRSPSRPPSPS HPSAHGSPAPVSTMTPKMSSEGPPRMSYKQRIHRHRSVAGRMSGGLFVSHNPP SEAAAPVARTSPAGTWSVSVGVRPLSPKTRHPRPROSSI GNSPSGVLASPOAG IIPAEAVMPVPAASPTPASPNRALTPTSLEAKDSLQDORONS PAGESKENVKASET SPFSKADNKGMSVSEHRKQIDDLKFRNDFRLOPSSSTGESMDQLLSKNREKSR DLIKDTAEAKOSFIDSSSSSSNCTSGSKTNSPSISPMSLNAEHRGPEVTSQGV QTSPPACKKEDREKDDTEQVRKSTLNPNAKEFNPRSPFSQPKSTPTTSPRPAQ PPSMVGHQAPAPVYTVQVCFAPNMYPVPVPGVQPLYPI PMTPMPVNOAKTYRAG VPNNQORQDQHOS TMWHPASAGPPIVATPAYSTOYVAYSPOQFPNOLVQHPH YQSHPHVYSPVIOGNARMWAPPAHAOPGLVSSAAOFGAEOHTAMVACPKLVNKE TSPSFFAITSGLAQOYAHNAAALHPTHPQPSATPTGQQSQHGGSHAPSPVQH HQHQALHLASQQOSAIYHAGLAPTPSPMTPASNTQSPSSFPAAQTFTVTHPS HVQPAYTTPHMAHVPOAHVQSGMVPSSHPTAHAPMMLMTQPPGPKAALAQSLQPIP VSTTAHPFPMYTHPSVQAHHOQOL"	
	ORIGIN		
	Query Match	70.0%; Score 3138.2; DB 9; Length 4225;	
	Best Local Similarity	86.2%; Pred. No. 0;	
	Matches 3662; Conservative	1; Mismatches 473; Indels 112; Gaps 14;	
QY	222	CGCGCAGCAGGTGGCCCGGCTGCGCTCGCTCAGCGCGCGCGGCGAGCGGCG	281
Db	71	CCCCAGCCGCGCCCGCGGGCTGGCGCGCGGCGCTCGCTCCCGCAGCGCGC	130
QY	282	GGCGCGGCGGTGGCGCGGCCCGCGGACCGTATCCCTCCGCCCGCCCTCC	341
Db	131	CGCGCGGCGGCGGCGCGGCGGTGCGGTATCCCTCCGCCCGCCCTCC	190
QY	342	CGGCCCGCTCCCTCCCGCAGAGCTCGCTCCCTCCGCCCTCAGACTGTTT	401
Db	191	CGGCGCCCTCCCGCGCGCGCGCTCGCCACCSTGCGCCTCAGACTGTTT	250
QY	402	CGGCAACGGCGCGCGGCTTTCGCGCGCGCTCCCGCGGCTCTTGTCTCG	461
Db	251	CGGCA-----CGCGGCTCCCGCGCGGCTCCCGCGGCTCGTCTCGG	304
QY	462	TCCCGCGCCCTTCGTCGTCGTCCTTCTCCCTCCGCCAGCCCGCGCC	521
Db	305	TCCCGCGCCCTTCGTCGTCGTCCTGCTGCTGCTG-----GCCCG	352
QY	522	GCCAAACCGCGCCTCCCGCTCGCGCGCGCTCGCTCCCGCGGCTTCCT	581



QY	2730	CAAGCGAATAGCCACGATTTTCCCTTCAATACTTAGTAAACAGGAGCAAGAGGGG	2789
Db	2510	CAAGCAACAGCCCTAGCATCTCCCTTCCATGCTTAGTAAATGACAGACCAAGAGGGG	2569
QY	2790	ACCTGAGGTCACTTCCCAAGGGGTTACAGCTTCCAGCCAGCATGTAACAGAGAAAGA	2849
Db	2570	GCTTGAAGTCAATCCCAAGGGGTGCAGCTTCCAGCCAGCGCTCAACCAAGAAAGGA	2629
QY	2850	CGATAAGGAAGAAAGAGACGACGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAA	2909
Db	2630	TGACAGAGAGAGAAAGACACACAGAGAGGTAGGAAATCGATTGAATCCCAA	2689
QY	2910	TGCAAGGAGTTCAACCCAGGTTCTTCTCTCAGCCAAAGGCTTCTATACCCCAATTC	2969
Db	2690	TGCAAGGAGTTCAACCCCTCGTTCTTCTCTCAGCCAAAGGCTTCTATACCCCAACGTC	2749
QY	2970	ACCTCGGCTCAAGCACAACTAGCCCATCTATGTGGGTATCAACAGCCAACTCCAGT	3029
Db	2750	ACCTCGGCTCAAGCACAACTAGCCCATCTATGTGGGTATCAAGCAGCGAGCTCCAGT	2809
QY	3030	TTATATCTCAGCCTGTTTGTGACCAAAATATGATGTATCCAGTCCCAGTGAGCCAGG	3089
Db	2810	GTACACTCAGCCTGTGCTTGGCACCCCAATATGATGTATCCGTTCCAGTGAGCCCGG	2869
QY	3090	CGTGCAACTTTATACCAATACCTATGAGGCCATGCGCAATGGAATCAAGCCAAACATA	3149
Db	2870	CGTACAACCTTTATACCAATACCTATGAGGCCATGCGTGTGAACCAAGCCAAACATA	2929
QY	3150	TAGAGC-----AGTACCATAATATGCCCAACAGCGGCAAGACAGCATCATCAGTGC	3203
Db	2930	TAGACAGGTAAGTACCAAAATATGCCCAACAGCGCAAGACCAACATCATCAAGCAC	2989
QY	3204	CATGATGCAACCCAGCGTCAGCAGCGGGCCACCGATTCGACGACCCCAACAGCTTACTC	3263
Db	2990	CATGATGCAACCCAGCGTCGCGGCAAGGCCACCCATCGTAGCCACCCCGCGCTTACTC	3049
QY	3264	CACGCAATATGTGCTACAGTCTCAGAGTTCCCAATCAGCCCTTGTTCAGCATGT	3323
Db	3050	CACCTCAGTACGTGCTACAGCCCTCAGCAGTTTCCCAATCAGCCCTTGTTCAGCATGT	3109
QY	3324	GCCATTTATCAGTCTCAGCATCTCATGTCTATAGTCTGTATATACAGGGTATGCTAG	3383
Db	3110	GCCGCAATTATCAGTCTCAGCATCTCATGTGTACAGTCTGTGTATACAGGTAATGCCAG	3169
QY	3384	AATGATGCGACCAACACACAGCCCGCTGGTTAGTATCTTCTTCAGCAACTCAGTA	3443
Db	3170	GATGATGCGACCAACAGCACATGCTCAGCCTGGTTAGTGTCTTCTTCAGTGTCTAGTT	3229
QY	3444	CGGGCTCATGACGACGCGATGCGATGTATGATGTATGTCATGTCCTCCAAATPACATACAAAGGA	3503
Db	3230	CGGGCTCACGACGACGCGACGCCATGTATGATGTATGTCATGTCCCAATPACATACAAAGGA	3289
QY	3504	GACAGCCCTTCTTCTACTTTGCAATTTCCAGGGCTCCCTTGTCTCAGCATATGCCA	3563
Db	3290	GACAGCCCTTCTTCTACTTTGCAATTTCCACGGCTCCCTTGTCTCAGCATATGCCA	3349
QY	3564	CCCTAACGCTACCTTCGACCCACATACCTCCACACCTTCAGCTTCAGCTACCCCACTGG	3623
Db	3350	TCCTAATGGCCCTTGCAATCCACATACCTCCCATCTCAGCTTCGGGCATCTCCACCGG	3409
QY	3624	ACAGCAGAAAGCCAAATGTGTGAAGTATCTCTGCAACCCAGTCTCTGTTCAGCACCATCA	3683
Db	3410	ACAGCAGAAAGCCAGCATGTGTGAAGTCAACCTGCACCCAGTCTCTGTTCAGCACCATCA	3469
QY	3684	GCACAGCCCGCCAGGCTCTCCATCTGGCCAGTCCACAGAGCATCAGCCATTACCA	3743
Db	3470	GCACAGGCTGCCAGGCTCTTCATCTGGCCAGTCCACAGAGCATCAGCCATTATCA	3529
QY	3744	CGCGGGCTTGGCCAACTCCACCTCCATGACACTGCTCTCAACAGCGAGTGCACACA	3803
Db	3530	TGCGGGCTTGGACCAACACCACTTCCATGACACTTGCTCTAATACAGACTCTCCACA	3589
QY	3804	GAATAGTTTCCAGCAGCACACAGACTGTCTTTACAGATCCATCTCTCAGTTTCAGCC	3863

Db	3590																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												</
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----







||||| 2272 GTTAAACCTCGTTCTTTCTCAGCAAGCCTTCTACTACCCCAAGCTCACTTCGGCC 2331  
||||| 2279 TCAAGCACAACCTAGCCCATCTATGGTGGGTGTCATCAACAGAGCAACTCCAGTTTATCTCA 3038  
||||| 2332 TCAAGCACAACCCAGCCCATCTATGGTGGGTGTCATCAGAGCCAGCTCCAGTGTCACCTCA 2391  
||||| 3039 GCCTGTTGTTTGGACCAAAATATGATGTATCCAGTCCAGTCCAGAGCCAGCGGTGCAACC 3098  
||||| 2392 GCCTGTGTGCTTGGCACCAAAATATGATGTATCCCGTCCAGTGAGCCCGGGCGTACAACC 2451  
||||| 3099 TTTATACCAATACCTATGAGGCCCATGCCAGTGAATCAAGCCAGAGCATATAGAC--- 3155  
||||| 2452 TTTATACCAATACCTATGAGGCCCATGCCAGTGAATCAAGCCAGAGCATATAGAGCAGG 2511  
||||| 3156 ---AGTACCAATATGCCCAACAGCGGCAAGACCAGAGCATATCAGAGTGCATGATGCA 3212  
||||| 2512 TAAAGTACCAATATGCCCAACAGCGAGAGACCAACATCATCAAGACCATGATGCA 2571  
||||| 3213 CCCAGCGTCAGAGCGGGCCCAACCGATTGCAAGCCACCCACAGCTTACTCCAGCAATA 3272  
||||| 2572 CCCAGCCTCCGCGGAGGGCCACCCCATCGTAGCCACCCCGCGCTTACTCCACTCAGTA 2631  
||||| 3273 TGTTCCTACAGTCTCAGCAGTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATTA 3332  
||||| 2632 CGTTGCCCTACAGCCCTCAGCAGTTTCCCAATCAGCCCTTTGGTCCAGCATGTGCCGCAATTA 2691  
||||| 3333 TCAGTCTCAGCATCTCATGTCTATAGTCCGTGTAATACAGGGTAAATGCTAGAAATGATGGC 3392  
||||| 2692 TCAGTCTCAGCATCTCATGTGTACAGTCCGTGTATACAGGTAATGCGAGATGATGGC 2751  
||||| 3393 ACCACCAACACAGCCCGCCAGCCTGGTTTGTAGTATCTTTTCAGCAACTCAGTACGGGGTCA 3452  
||||| 2752 ACCACGACACATGCTCAGCCTGGTTTGTAGTGTCTTTTCAGTGTCTCAGTTTCGGGGCTCA 2811  
||||| 3453 TGAGCAGAGCATGCGATGTATGATGTCCCAAAATACCATACACAGAGGACAGGCC 3512  
||||| 2812 CGAGCAGAGCGACGGCATGTATGATGTGCCAAATTTACCATACACAGAGGACAGGCC 2871  
||||| 3513 TTCTTTTCTACTTTGTCATTTCCACGGGCTCCCTGTCTCAGCAGATGTGGCCACCTTAAGCC 3572  
||||| 2872 TTCTTTTCTACTTTGTCATTTCCACGGGCTCCCTGTCTCAGCAGATGTACATCTCTATGC 2931  
||||| 3573 TACCTTGACCCACATATCTCCACACCTTAGCCTTTACGTACCTACCCCATCTGACAGCAGCA 3632  
||||| 2932 CGCCTTGATCCACATACTCCCATCTCAGCCTTTCCGGCCACTCCCAACCGGACAGCAGCA 2991  
||||| 3633 AAGCCAAACATGTTGAAGTTCATCTCGACCCAGTCCCTGTTCAGCACCATCAGCCAGCCAGC 3692  
||||| 2992 AAGCAGCATGTTGAAGTTCACCTTGCAACCCAGTCCCTGTTCAGCACCATCAGCCAGCCAGC 3051  
||||| 3693 CGCCAGGCTCTCCATCTCGGCAGTCCACAGCAGCAGTTCAGCCATTTTACCAACCGGGGCT 3752  
||||| 3052 TGCCAGGCTCTTTCATCTCGGCAGTCCACAGCAGCAGTTCGGCCATTTTATCATCGGGGCT 3111  
||||| 3753 TGGGCCAACTCCACTCCATGACACCTTCCTCCAAACGCGAGTCCGACAGAAATAGTTT 3812  
||||| 3112 GGCACCAACACCCCTTCATGACACCTTCCTCTAATACACAGTCTCCACAGAGCAGTTT 3171  
||||| 3813 CCCAGCAGCAGACAGACTGTCTTTACGATCATCTCTTCTCAGTTCAGCCGGGCTATAC 3872  
||||| 3172 CCCAGCAGCAGCAACAGACAGTCTTCCACCATCACCTTCTCATGTTTCAGCCGGCATAC 3231  
||||| 3873 CAACCCACCCACATGCGCCACGTACCTCAGGCTCATGTACAGTCAAGGAATGGTTCTTC 3932  
||||| 3232 CACCCCAACCCACATGCGCCACGTACTCTCAGGCTCATGTACAGTCAAGGAATGGTTCTTC 3291  
||||| 3933 TCATCCAACTGCGCCATGCGCCAAATGATGCTAATACGACAGCCACCCGCGGTCCCA 3992  
||||| 3292 TCATCCAACTGCGCCATGCGCCAAATGATGCTAATACGACAGCCACCC---CGGTCCCA 3348  
||||| 3993 GCGCGCCTCGCTCAAGTGACATACAGCCCATTCAGTCTCGACAAACAGCGCATTTCCC 4052  
|||||

Db 3349 GGCGCCCTCGCTCAAGTGCACTACAGCCCATTTCCAGTTTCGACAAACAGGCGCATTTCCC 3408  
Qy 4053 CTATATGACGACCCCTTTCAGTACAAAGCCCAACCAACAGCAGTGTGAAGCTGCGCTGG 4112  
Db 3409 TTATATGACGACCCCTTTCAGTACAAAGCCCAACCAACAGCAGTGTGAAGCTGCGCTGG 3468  
Qy 4113 AGAACCGAAAGGCGCAATTCCTCCTCTTCTACTGCTTCTACCAACTCGAAGCAGACAG 4172  
Db 3469 AGAACCGAAAGGCGCAATTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3527  
Qy 4173 AAAACTAGAAATTCATTTATTTTGTATTTTAAATATATATATGTTGATTTCTTGTAACTCC 4232  
Db 3528 AAAACTAGAAATTCATTTATTTTGTATTTT-----TAAAGATACACTGATTTAACTCT 3581  
Qy 4233 AATAGGAATGCTAAACAGTTCACCTTGCAGTGAAGAT-ACITTTGGACCGAGTAGAGGCATTT 4291  
Db 3582 GATAGGAATGCTAAACAGTTCACCTTGCAGTGAAGATGTTTGGACCGAGTAGAGGCATGT 3641  
Qy 4292 AGAACCTTGGGGGCTATTCATTAATTCATATGCTGTTTCAGAGTCCCGCAGGTAACCCCA 4351  
Db 3642 AGGACTT-GTGGCTGTTCATAATTCATGTGCTTTCAGAGGTCCTCTGCAAGTA-CCCA 3699  
Qy 4352 GCTCTGCTTCCGAAACCTGGAAGTTATTTTAAATACCCCTTGAAGTTCATGAACA 4411  
Db 3700 GCTCTGCTTCTGAAACTGGAAGTTATTTTAAATGGCCCTTGAGAGTTCATGAACA 3759  
Qy 4412 CATCAGCTAGCAAAAGAAAGTAACAAGAGTGAATCTTCTGCT 4450  
Db 3760 CATCAGCTAGCAAAAGAAAGTAACAAGAGTGAATCTTCTGCT 3798

RESULT 12  
AK095017  
LOCUS AK095017  
DEFINITION Homo sapiens cDNA FLJ37698 fis, clone BRHIP2015679, highly similar to Human ataxin-2 (SCA2) mRNA.  
ACCESSION AK095017  
VERSION AK095017.1 GI:21754198  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1  
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Tanaka, H., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirao, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togawa, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terahashi, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Taniguchi, A., Fujiwara, T., Ono, T., Yanada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,



DEFINITION	Xenopus laevis cDNA clone MGC:115230 IMAGE:5079818, complete cds.
ACCESSION	BC097692
VERSION	BC097692.1 GI:66910767
KEYWORDS	MGC.
SOURCE	Xenopus laevis (African clawed frog)
ORGANISM	Xenopus laevis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
REFERENCE	Xenopodinae; Xenopus; Xenopus.
AUTHORS	1. (bases 1 to 2558)
	Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
	and Richardson,P.
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative
JOURNAL	Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED	12454917
REFERENCE	2. (bases 1 to 2558)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uesdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
CONSRTM	Mammalian Gene Collection Program Team
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	3. (bases 1 to 2558)
AUTHORS	Klein,S. and Gerhardt,D.S.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUN-2005) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
REMARK	NIH-MGC Project
COMMENT	Contact: XGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a> Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	source
	Location/Qualifiers
	1..2558
	/organism="Xenopus laevis"
	/mol_type="mRNA"
	/db_xref="taxon:8355"
	/clone="MGC:115230 IMAGE:5079818"
	/tissue_type="Oocytes"

ORIGIN

Query Match

Best Local Similarity

Matches 1504; Conservative

73.1%; Pred. No. 1.6e-150;

0; Mismatches 492; Indels 61; Gaps 10;

23.4%; Score 1047.8; DB 5; Length 2558;

QY

863

CCGCGCGGCGGAGCCCGCCCTGGCAGAGGTGCGAAACAGTAAACAAAGGACTGCTCTCAGT

922

DB

301

CAGGAGGAGGAAGGAGCAAGATATGGCAGAGGAGGATCAGTGGAAAGGGCCCCCTCAAT

360

QY

923

CTACGATTTCTTTTGTATGGAATCTATGCAATATATGAGGATGGTTTCATATATTACATCAG

982

DB

361

---CGATCTCGTTTGTATGTTATGCAATATATGAGACTGTTTCATATTTTACATCAG

417

QY

983

TTGTTGGTCCCAATGTGAAGTCAAGTGAAATATGGAGGTATATATGAAGGAGTTTTTA

1042

DB

418

TAGTTGGTCCAAATGTGAGGTATTTGGTGAATAATGGAAACGTATATGAAGGTGTTTTTA

477

QY

1043

AAACTTACAGTCCGAAGTGTGATTTGGTACTTTGATGCGGCACATCAGAAAAAGTACAGAAAT

1102

DB

478

AACTTACAGTCCCTAAGTGGCATTTGGTGCTTGTGATGAGCTCATAAGAAAACTACAGAAAT

537

QY

1103

CCAGTTGGGGCCGAAACGCTGAAGAAATAATGAGAGATATTTTGTTCAAATGTTTCAGACT

1162

DB

538

CTATTGGGGCCCAAGCGTGAAGACATTGTGGATAGTATATTTGTTAAATCTTCCGACT

597

QY

1163

TTGTTGGTGCAGTTTAAAGATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTG

1222

DB

598

TTGTTTGGTGCAGTTTAAAGATATGGATGTGAATTTATGCAAGACGAGATGGGTTTCAG

657

QY

1223

ACTCTGTCTATCAGTCTAAAGTGAATGGCGAACACAAAGAGAGGACCTGGAGCCCTGGG

1282

DB

658

ATTCCTGGAATGCTTGTAGGTTAATGTTGTAACATAGGAGAGAGATCTGGAGCCGTGG

717

QY

1283

ATGCAGGTGA-----ACTCAGCCCAATGAGAGAACTTGAGGCTTTGG

1324

DB

718

ATGGTGGTGAAGTCTCATTTAGCAGCACACCTACTTGGCAAGATGAGCTGGAATCTTTGG

777

QY

1325

AAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATAATCAAGAAAAATT

1384

DB

778

AAATGA---GTCTAATGGATGGGATCCAAATGAATGTTTTCGCTACATGAAGAAAAATT

834

QY

1385

ATGGTGTAGTGTCTACGTATGATGACAGTTTATCTTCGTATACAGTGCCTTAGAGAG

1444

DB

835

ATGGTGTGTGTCAACTTATGACAGTAGTCTGTCTCTGTCACACTGTTCCATTAGAAAGG

894

QY

1445

ATACTCAGAAAGAAATTTTAAAAACGGAGAGCAAGGCAACACCATGTAGCAGAGAAATTG

1504

DB

895

ATAATTCAGAAAGAAATTTTGAAGAGAGAGAGCAAGCAGACGAGCTCAGATTGCTGAG

954

QY

1505

AGTCAAGTGCCAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATGATAGGAGTAGGAG

1564

DB

955

AATCAAGTTCCTCAGTATAAAGCTAGATTGCTTTGGAAAAATGATGAGCGATCAGAGGAG

1014

CDs

/clone\_lib="NICHD\_XGC\_001"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

228..2120

/codon\_start=1

/product="Unknown (protein for MGC:115230)"

/protein\_id="AAH97692.1"

/db\_xref="GI:66910768"

/translation="MSMKQAQANNRKPGGSVPAAGGAGGAGGGRNMGRGRISGKGPPO

ESIYVGRKEDIVDSILFKSSDFVMVQKMDVNYARRDGTDSGIACPLNHEHKKTT

EPWGGSLSSITGTGDELSLENSNGWDPNEMFRYNEENYGVSTYDLSLSYTV

RHSLNRENYKIPVQNKDAMLTHTHSQRNSARIGQSGSPITPSRTGSHSSDYNPN

AGADRVVGGVPMPCSP

SHPSAGSPAPVMPKRMSSGPPRSPKAPRTHRVSNRGTGTCGGLFVSRNE

AGEASVQPVARNSSGGTWSVSVQVRLSPKTHRPSRPLGSHAPGSNATSPQGTV

PFESVMPVPAASPTPASPADRAVTSCEAKOTRLQDQRONSPAGCRENSKYSESSK

TENKAVSPMASEQRKQDDLDLKKFKNDFRLQFSSPDLDLPVKRSVEKPRDPLKK

KNGFKC"



[illegible]

Db 899 CAGAGAACTGCAGTCACCGGAGGGGCATGGCCCCAACACTAGGGACAAATAATATATT 958  
QY 1642 CCTCTTGGACAAAGAAATAGAGAAGTCATATCTTGGGGAAGTGGGAGACAGAAATTCACCG 1701  
Db 959 CCTCTTGGACAAAGAAACAGAGAAGTCCTATCTTGGGGAAGTGGGAGACAGAGCTCACCA 1018  
QY 1702 CGTATGGGCCAGCCTGGATCGGGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGAT 1761  
Db 1019 CGGATGGGCCAGCCTGGCCAGGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGAT 1078  
QY 1762 TTCAACCCGGAATTCCTGGTTTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCG 1821  
Db 1079 TTCAACCCGGAACGCTGGCTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCG 1138  
QY 1822 CCTTGGCCATCTCCTTCTCTCGCCACCTTCTCGCTACAGTCAGGTCCCAACTCTCTT 1881  
Db 1139 CCTTGGCCATCTCCTTCTCTCGCCACCTTCTCGCTACAGTCAGGTCCCAACTCTCTT 1198  
QY 1882 CCACCTCGGGCAGCCACCCCTACACGGCGCCCTCC 1917  
Db 1199 CCACCTCGGGCAGCCACCCCTACACGGCCTCGTGCC 1234

Search completed: December 23, 2005, 00:44:05  
Job time : 21615 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 15:52:25 ; Search time 2343 Seconds  
(without alignments)

12746.260 Million cell updates/sec

Title: US-10-802-228-1

Perfect score: 4481

Sequence: 1 acccccgaagcaacccca.....taaaaaaaaaaaaaaaaaaaaa 4481

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002as.\*

7: Geneseqn2002bs.\*

8: Geneseqn2003as.\*

9: Geneseqn2003bs.\*

10: Geneseqn2003cs.\*

11: Geneseqn2003ds.\*

12: Geneseqn2004as.\*

13: Geneseqn2004bs.\*

14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4481	100.0	4481	2	AAV06552 Human SCA
2	4481	100.0	4481	2	Aaz23428 Human SCA
3	4481	100.0	4481	10	ADD18753 Human dis
4	4481	100.0	4481	10	Adh69466 Human spi
5	4481	100.0	4481	10	Abv75291 Human SCA
6	4481	100.0	4481	14	Adz49164 Insulin s
7	4347.8	97.0	4367	2	Aav30270 Gene caus
8	3949	88.1	4200	2	Aat78912 Spinocere
9	3138.2	70.0	4225	10	Abv75290 Mouse SCA
10	2957	66.0	3798	2	Aav06553 Mouse SCA
11	1730.6	38.6	2006	3	Aaf21628 Human bre
12	935.2	20.9	1257	10	Abv75289 Mouse SCA
13	633	14.1	4369	13	Adr08336 Full leng
14	571.2	12.7	623	2	Aav17229 SCA2 gene
15	497.4	11.1	516	2	Aav06551 SCA2 gene
16	455.8	10.2	459	10	Adh69479 Human SCA
17	441.2	9.8	704	6	Abq36595 Oligonuc1
18	441.2	9.8	704	6	Abq36594 Oligonuc1
19	437.6	9.8	485	9	Ach41056 Human foe

c	20	412.4	9.2	73995	11	ACN43986	Acn43986 Human gen
	21	354.2	7.9	355	2	AAV17224	Aav17224 SCA2 gene
	22	301	6.7	499	13	ADQ51680	Adq51680 Novel can
c	23	288.2	6.4	93631	11	ACN43984	Continuation (5 of
	24	274.2	6.1	463	6	ABV87938	Abv87938 Human col
	25	271.4	6.1	704	6	ABQ36596	Abq36596 Oligonuc1
c	26	271.4	6.1	704	6	ABQ36597	Abq36597 Oligonuc1
	27	268	6.0	4674	2	AAx90421	Aax90421 Human ata
	28	268	6.0	4674	8	ABZ72015	Abz72015 Human ata
	29	268	6.0	4674	10	ADH29078	Adh29078 Human chr
c	30	266.4	5.9	1707	6	ABK34831	Abk34831 Human chr
c	31	243.2	5.4	346	5	AAx79688	Aax79688 DNA encod
c	32	243.2	5.4	346	5	AAx91261	Aax91261 DNA encod
c	33	243.2	5.4	370	5	AAx91972	Aax91972 DNA encod
c	34	243.2	5.4	370	5	AAx87785	Aax87785 DNA encod
	35	240	5.4	1226	6	ABK84413	Abk84413 Human CDN
	36	222.8	5.0	392	10	ADD33138	Add33138 Human mit
	37	214.6	4.8	10378	4	ABK43051	Abk43051 Genomic s
	38	214.6	4.8	10378	9	ADB61207	Adb61207 Connectiv
c	39	195.4	4.4	278	10	ADD33137	Add33137 Human mit
c	40	187.8	4.2	320	4	AAx58443	Aax58443 cDNA #111
	41	183	4.1	183	10	ADH69482	Adh69482 Human ata
	42	177	4.0	177	10	ADH69483	Adh69483 Human ata
	43	169	3.8	169	10	ADH69481	Adh69481 Human ata
	44	167	3.7	205	2	AAV17225	Aav17225 SCA2 gene
	45	131.2	2.9	470	4	ABA57581	Abas7581 Human foe

## ALIGNMENTS

RESULT 1	AAV06552	AAV06552 standard; cDNA; 4481 BP.
ID	AAV06552	standard; cDNA; 4481 BP.
XX	AAV06552;	
AC	AAV06552;	
XX	06-JUL-1998	(first entry)
DT	06-JUL-1998	(first entry)
DE	Human SCA2 cDNA including CAG repeat region.	
KW	SCA2 gene; spinocerebellar ataxia-2; ataxin-2; human; diagnosis;	
KX	olivo-ponto-cerebellar atrophy; ss; ds.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	164..4101
FT	primer_bind	/tag= a
FT	repeat_region	complement(631..648)
FT	repeat_unit	/note= "primer SCA2-A binding site"
FT	repeat_unit	658..723
FT	repeat_unit	/tag= f
FT	repeat_unit	/note= "CAG repeat region"
FT	repeat_unit	658..660
FT	repeat_unit	/tag= g
FT	repeat_unit	/note= "CAG repeat"
FT	repeat_unit	661..663
FT	repeat_unit	/tag= h
FT	repeat_unit	/note= "CAG repeat"
FT	repeat_unit	664..666
FT	repeat_unit	/tag= i
FT	repeat_unit	/note= "CAG repeat"
FT	repeat_unit	667..669
FT	repeat_unit	/tag= j
FT	repeat_unit	/note= "CAG repeat"
FT	repeat_unit	670..672
FT	repeat_unit	/tag= k
FT	repeat_unit	/note= "CAG repeat"
FT	repeat_unit	673..675
FT	repeat_unit	/tag= l
FT	repeat_unit	/note= "CAG repeat"

FT	repeat_unit	676. .678	PS	Claim 6; Page 52-58; 98pp; English.
FT		/tag= m	XX	This cDNA sequence corresponds to a novel SCA2 gene encoding a human
FT	repeat_unit	/note= "CAG repeat"	CC	spinocerebellar ataxia-2 (SCA2) polypeptide, designated ataxin-2 (see
FT		679. .681	CC	AAW33807). A trisomy 21 foetal brain cDNA library and an adult human
FT	repeat_unit	/tag= n	CC	frontal cortex cDNA library in lambda ZapII were screened with probes
FT		/note= "CAG repeat"	CC	obtained by PCR amplification of plasmid AAP65122B (see AAV06551). PCR
FT		685. .687	CC	products were used to screen the human adult frontal cortex library, and
FT	repeat_unit	/tag= o	CC	5' clones were obtained by RT-PCR of placental mRNAs. Overlapping clones
FT		/note= "CAG repeat"	CC	was used to generate the composite 4481 bp sequence. Ataxia type 2 can be
FT		688. .690	CC	diagnosed by detecting a genomic or transcribed mRNA sequence in an
FT	repeat_unit	/tag= p	CC	individual having an expanded CAG repeat at a location corresponding to
FT		/note= "CAG repeat"	CC	the CAG repeat region of the SCA2 gene. The presence of at least 13 CAG
FT		691. .693	CC	repeats above the normal level (22, occasionally 23, repeats) is
FT	repeat_unit	/tag= q	CC	indicative of SCA2. Primers (see AAT99640-41) amplifying at least this
FT		/note= "CAG repeat"	CC	region are used for diagnosis. Also claimed are kits for detecting
FT		694. .696	CC	mutations at the SCA2 locus, antisense oligonucleotides, and transgenic
FT	repeat_unit	/tag= r	CC	animals useful for studying the physiological roles of ataxin-2 and its
FT		/note= "CAG repeat"	CC	effect upon behaviour
FT		700. .702	XX	
FT	repeat_unit	/tag= s	SQ	Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
FT		/note= "CAG repeat"		Query Match 100.0%; Score 4481; DB 2; Length 4481;
FT	repeat_unit	703. .705		Best Local Similarity 100.0%; Pred. No. 0;
FT		/tag= t		Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT	repeat_unit	/note= "CAG repeat"		
FT		706. .708		
FT		/tag= u		
FT	repeat_unit	/note= "CAG repeat"	Qy	1 ACCCCGAGAAAGCAACCCAGCGCGCCGCTCCTCAGGTGTCCCTCCCGGCCCGGG 60
FT		709. .711	Db	1 ACCCCGAGAAAGCAACCCAGCGCGCCGCTCCTCAGGTGTCCCTCCCGGCCCGGG 60
FT	repeat_unit	/tag= v		
FT		/note= "CAG repeat"		
FT		712. .714		
FT	repeat_unit	/tag= w	Qy	61 GCCACCTCAGTTCTGCTTCGGTGTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
FT		/note= "CAG repeat"	Db	61 GCCACCTCAGTTCTGCTTCGGTGTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
FT		715. .717		
FT	repeat_unit	/tag= x	Qy	121 CACCTCCGCTCCACCCCGGCGCTCGGGCGCGCCCTCGATGCGCTCAGCGGGCGCA 180
FT		/note= "CAG repeat"	Db	121 CACCTCCGCTCCACCCCGGCGCTCGGGCGCGCCCTCGATGCGCTCAGCGGGCGCA 180
FT		718. .720		
FT	repeat_unit	/tag= y	Qy	181 GTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCCGCGAGCCAGTGGCCC 240
FT		/note= "CAG repeat"	Db	181 GTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCCGCGAGCCAGTGGCCC 240
FT		721. .723		
FT	repeat_unit	/tag= z	Qy	241 GGGTGGCGTCTGCTTCAGCGCGCGCGCGCGAGCGGGCGGGCGCGGTGGCGCGCC 300
FT	primer_bind	/note= "CAG repeat"	Db	241 GGGTGGCGTCTGCTTCAGCGCGCGCGCGCGAGCGGGCGGGCGCGGTGGCGCGCC 300
FT		740. .757		
FT	exon	/tag= c	Qy	301 CCGGAGCGGTATCCCTCGCGCGCCCTCCCGCGCGCGCGCGCGCCCTCCCTCCCGG 360
FT		/note= "primer SCA2-B binding site"	Db	301 CCGGAGCGGTATCCCTCGCGCGCCCTCCCGCGCGCGCGCGCGCCCTCCCTCCCGG 360
FT		899. .900		
FT	primer_bind	/tag= e	Qy	361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCAACGGCGGGCGCGC 420
FT		/note= "predicted splice site"	Db	361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCAACGGCGGGCGCGC 420
FT		1070. .1091		
FT		/tag= d	Qy	421 TTTCGGCCCCGGCTCCCGCGGCTCTTTGGTCTCGCGGGCTCTCCCGCCCTTCGTCTC 480
FT		/note= "primer SCA2-14B binding site"	Db	421 TTTCGGCCCCGGCTCCCGCGGCTCTTTGGTCTCGCGGGCTCTCCCGCCCTTCGTCTC 480
FT				
FT			Qy	481 GTCTTCTCCCTTCGCGAGCCGGCGCCCTCCGCGCGGGCCCAACCGGGCCCTCCCGC 540
FT			Db	481 GTCTTCTCCCTTCGCGAGCCGGCGCCCTCCGCGCGGGCCCAACCGGGCCCTCCCGC 540
FT			Qy	541 CTGGCGCCCGTGTGCTCCCGCGCGCTTCGGGTCTCGGGCGGTCTCTTTGGCGCCCGG 600
FT			Db	541 CTGGCGCCCGTGTGCTCCCGCGCGCTTCGGGTCTCGGGCGGTCTCTTTGGCGCCCGG 600
FT			Qy	601 TGTCCCGCCCGCGTTCGAGCCCGGTGTATGGGCCCTCTCAACATGTCTCGCTGAAGCC 660
FT			Db	601 TGTCCCGCCCGCGTTCGAGCCCGGTGTATGGGCCCTCTCAACATGTCTCGCTGAAGCC 660
FT			Qy	661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
FT			Db	661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720

NUC1  
Nucleic acids encoding human and mouse ataxin 2 - a product of the  
spinocerebellar ataxia 2 gene, SCA2; useful in the diagnosis of ataxia  
type 2.

Qy	721	CAGCCGCCCGCGGCTGCCAATGTCCGCAAGCCGGCGGCGAGCGGCTTCTAGCGTCG	780
Db	721	CAGCCGCCCGCGGCTGCCAATGTCCGCAAGCCGGCGGCGAGCGGCTTCTAGCGTCG	780
Qy	781	CCGCGCGCGGCGCTTCGCGCTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGGCGCGGCTCCC	840
Db	781	CCGCGCGCGGCGCTTCGCGCTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGGCGCGAGCTCCC	840
Qy	841	TCCTCGGTGGTTCGCGGCGACTCCGCGCGGGGAGGCCCGGCTTGGGCGAGAGCTGCAAAAC	900
Db	841	TCCTCGGTGGTTCGCGGCGACTCCGCGCGGGGAGGCCCGGCTTGGGCGAGAGCTCGAAGC	900
Qy	901	AGTAACAAAGGACTGCGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATATGAGG	960
Db	901	AGTAACAAAGGACTGCGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATATGAGG	960
Qy	961	ATGGTTTCATATACATACATCAGTTGTTGGCTCCAAATGTGAAGTACAAAGTGAAAAATGGA	1020
Db	961	ATGGTTTCATATACATACATCAGTTGTTGGCTCCAAATGTGAAGTACAAAGTGAAAAATGGA	1020
Qy	1021	GGTATATATGAAGGAGTTTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC	1080
Db	1021	GGTATATATGAAGGAGTTTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC	1080
Qy	1081	GCACATGAGAAAAGTACAGAAATCAGTTTCGGGGCCGAAAACTGAGAGAAATAATGAGAGT	1140
Db	1081	GCACATGAGAAAAGTACAGAAATCAGTTTCGGGGCCGAAAACTGAGAGAAATAATGAGAGT	1140
Qy	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGGAATCCAGTTAT	1200
Db	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGGAATCCAGTTAT	1200
Qy	1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAAATGCGCAACACAAA	1260
Db	1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAAATGCGCAACACAAA	1260
Qy	1261	GAGAAGACCTGGAGCCCTGGGATGCAAGTGAATCTCAGGCCAATGAGGAACCTTGAGGCT	1320
Db	1261	GAGAAGACCTGGAGCCCTGGGATGCAAGTGAATCTCAGGCCAATGAGGAACCTTGAGGCT	1320
Qy	1321	TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATATGAAGAA	1380
Db	1321	TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATATGAAGAA	1380
Qy	1381	AAATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTGAGAA	1440
Db	1381	AAATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTGAGAA	1440
Qy	1441	AGAGTAACCTCAGAAGAAATTTTAAACGGGAAGCAAGGCAACACCACTAGTTAGCAGAGAA	1500
Db	1441	AGAGTAACCTCAGAAGAAATTTTAAACGGGAAGCAAGGCAACACCACTAGTTAGCAGAGAA	1500
Qy	1501	ATTGAGTCAAGTGCCAGTACAAAGCTCAGTGGCCCTTGGAAAAATGATAGGAGTGAG	1560
Db	1501	ATTGAGTCAAGTGCCAGTACAAAGCTCAGTGGCCCTTGGAAAAATGATAGGAGTGAG	1560
Qy	1561	GAAGAAAATAACACAGCAGTTTCCAGTGAACGTGAGGGGCGACAGCATTAAC	1620
Db	1561	GAAGAAAATAACACAGCAGTTTCCAGTGAACGTGAGGGGCGACAGCATTAAC	1620
Qy	1621	ACTAGGGAATAATAATATTTCTTCCTGGACAAAGAAATAGAGAGTCTATCTCTGGGA	1680
Db	1621	ACTAGGGAATAATAATATTTCTTCCTGGACAAAGAAATAGAGAGTCTATCTCTGGGA	1680
Qy	1681	AGTGGGAGACAGAAATTCACCGCGTATGGGCCAGCGCTGGATCGGGCTCCATGCCATCAAGA	1740
Db	1681	AGTGGGAGACAGAAATTCACCGCGTATGGGCCAGCGCTGGATCGGGCTCCATGCCATCAAGA	1740
Qy	1741	TCCACTTCTCACACTTCAGATTTCAAACCGAATTTCTGGTTTCAGACCAAGAGTAGTTAAT	1800
Db	1741	TCCACTTCTCACACTTCAGATTTCAAACCGAATTTCTGGTTTCAGACCAAGAGTAGTTAAT	1800

Qy	1801	GGAGGTGTTCCCTGGGCATCGCCTTGGCCCATCTCCTTCTCTCTCGCCACGTTTCTCGCTAC	1866
Db	1801	GGAGGTGTTCCCTGGGCATCGCCTTGGCCCATCTCCTTCTCTCTCGCCACGTTTCTCGCTAC	1860
Qy	1861	CAGTCAGGTCCCAACTCTCTTCCACCTCGGCAGGCACCCCTACACGGCGCCCTCCAGG	1920
Db	1861	CAGTCAGGTCCCAACTCTCTTCCACCTCGGCAGGCACCCCTACACGGCGCCCTCCAGG	1920
Qy	1921	CCCCCTCGGGCCATCCAGACCCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCT	1980
Db	1921	CCCCCTCGGGCCATCCAGACCCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCT	1980
Qy	1981	GTCTCTACTATGCTTAAACGCGATGTCTTCCAGAGGGCCTCCAGAGATGTCCCCAAAGGCC	2040
Db	1981	GTCTCTACTATGCTTAAACGCGATGTCTTCCAGAGGGCCTCCAGAGATGTCCCCAAAGGCC	2040
Qy	2041	CAGGCATCCTCGAATACACAGATTCTCTGCTGGAGGGTTCCATATCCAGTGGCCTA	2100
Db	2041	CAGGCATCCTCGAATACACAGATTCTCTGCTGGAGGGTTCCATATCCAGTGGCCTA	2100
Qy	2101	GAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGCACAGT	2160
Db	2101	GAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGCACAGT	2160
Qy	2161	CCCTCGGGGGAAAGTGGTTCATCAGTGGTTCAGTGGGTTCCAGATTTATCCCTAAAACT	2220
Db	2161	CCCTCGGGGGAAAGTGGTTCATCAGTGGTTCAGTGGGTTCCAGATTTATCCCTAAAACT	2220
Qy	2221	CATAGACCCAGGTCTCCACACAGACAGATTTGGAAATAACCCAGTGGGCCAGTTCTT	2280
Db	2221	CATAGACCCAGGTCTCCACACAGACAGATTTGGAAATAACCCAGTGGGCCAGTTCTT	2280
Qy	2281	GCTTCTCCCAAGCTGGTATTTATTCAACTGAAAGCTGTTGCCATGCTATTCCAGCTGCA	2340
Db	2281	GCTTCTCCCAAGCTGGTATTTATTCAACTGAAAGCTGTTGCCATGCTATTCCAGCTGCA	2340
Qy	2341	TCCTCTACGCTCTAGTCTCTGATCGAACAGAGCTGTTACCCCTTCTAGTAGGCTAAA	2400
Db	2341	TCCTCTACGCTCTAGTCTCTGATCGAACAGAGCTGTTACCCCTTCTAGTAGGCTAAA	2400
Qy	2401	GATTCCAGGCTTCAAGATCAGAGCGAAGCTCTCTCGACGGAAATAAGAAATATTAAA	2460
Db	2401	GATTCCAGGCTTCAAGATCAGAGCGAAGCTCTCTCGACGGAAATAAGAAATATTAAA	2460
Qy	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTTGAAACAAAGGTATATCACAGTTGTT	2520
Db	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTTGAAACAAAGGTATATCACAGTTGTT	2520
Qy	2521	TCTGAACATAGAAAAAGATTTGATTTAAAGAAATTTAAGAAATGATTTTAGTTACAG	2580
Db	2521	TCTGAACATAGAAAAAGATTTGATTTAAAGAAATTTAAGAAATGATTTTAGTTACAG	2580
Qy	2581	CCAAGTTCTACTTCTGAATCTATGGATCACTACTAAACAAAAATAGAGGGGAGAAAA	2640
Db	2581	CCAAGTTCTACTTCTGAATCTATGGATCACTACTAAACAAAAATAGAGGGGAGAAAA	2640
Qy	2641	TCAAGATTTGATCAAAAGCAAAATTTGAACCAAGTCTAAGGATCTTTTCAATGAAAT	2700
Db	2641	TCAAGATTTGATCAAAAGCAAAATTTGAACCAAGTCTAAGGATCTTTTCAATGAAAT	2700
Qy	2701	AGCAGCAGCAACTGTATACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCCCTCA	2760
Db	2701	AGCAGCAGCAACTGTATACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCCCTCA	2760
Qy	2761	ATACCTTAGTAACCGGAGCAAGAGGGGACCTGAGTCACTTCCCAGGGGTTTCAGACT	2820
Db	2761	ATACCTTAGTAACCGGAGCAAGAGGGGACCTGAGTCACTTCCCAGGGGTTTCAGACT	2820
Qy	2821	TCCAGCCCGCATGTAAACAGAGAAAGCGATTAAGGAAGAAAGACGCGACTGAG	2880
Db	2821	TCCAGCCCGCATGTAAACAGAGAAAGCGATTAAGGAAGAAAGACGCGACTGAG	2880
Qy	2881	CAAGTTAGGAAATCAAATTTGAATTCCCAATGCAAAAGGAGTTCAACCCACGTTCTCTCT	2940

||||| 2881 CRAATTAGAAATCAACATTTGATCCCAATGAAAGGAGTTCAACCCACGTTCTTCTCT 2940  
QY 2941 CAGCCAAAGCCCTTCTACTACCCCAACTTTCACCTCGGCTCAAGCACAACCTTAGCCCATCT 3000  
Db 2941 CAGCCAAAGCCCTTCTACTACCCCAACTTTCACCTCGGCTCAAGCACAACCTTAGCCCATCT 3000  
QY 3001 ATGGTGGTCAATCAACAGCCAACTCCAGTTTATATCTAGCCTGTTTGTGTTGTCACCAAT 3060  
Db 3001 ATGGTGGTCAATCAACAGCCAACTCCAGTTTATATCTAGCCTGTTTGTGTTGTCACCAAT 3060  
QY 3061 ATGATGATATCCAGTCCAGTCCAGCCAGCGTGCAACTTTATATACCCCAATACCTATGAG 3120  
Db 3061 ATGATGATATCCAGTCCAGTCCAGCCAGCGTGCAACTTTATATACCCCAATACCTATGAG 3120  
QY 3121 CCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAATATATGCCCAACAGCGG 3180  
Db 3121 CCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAATATATGCCCAACAGCGG 3180  
QY 3181 CAAGACAGCATCATCAGAGTGCATATGCAACCCAGCGTCAGCAGCGGGCCCAACCGATT 3240  
Db 3181 CAAGACAGCATCATCAGAGTGCATATGCAACCCAGCGTCAGCAGCGGGCCCAACCGATT 3240  
QY 3241 GCAGCACCACCCAGCCTTACTCCACGCAATATGTCCTACAGTCCCTCAGCAGTCCCA 3300  
Db 3241 GCAGCACCACCCAGCCTTACTCCACGCAATATGTCCTACAGTCCCTCAGCAGTCCCA 3300  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT 3360  
Db 3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT 3360  
QY 3361 CCTGTATACAGGTAATGCTAGAAATGATGCGACCAACAGCCGACCGCTGGTTTA 3420  
Db 3361 CCTGTATACAGGTAATGCTAGAAATGATGCGACCAACAGCCGACCGCTGGTTTA 3420  
QY 3421 GTATCTTCTTCAGCAACTCAGTAGGGGCTCATGAGCAGCGCATGCGATGATGCAATG 3480  
Db 3421 GTATCTTCTTCAGCAACTCAGTAGGGGCTCATGAGCAGCGCATGCGATGATGCAATG 3480  
QY 3481 CCCAAATTAACATCAACAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540  
Db 3481 CCCAAATTAACATCAACAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540  
QY 3541 TCCCTTGTCTAGCAGTATGCGGACCCCTAAACGCTACCTTGCACCCACATCTCCACCCCT 3600  
Db 3541 TCCCTTGTCTAGCAGTATGCGGACCCCTAAACGCTACCTTGCACCCACATCTCCACCCCT 3600  
QY 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCCAACATGTTGGAGTCATCTGCA 3660  
Db 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCCAACATGTTGGAGTCATCTGCA 3660  
QY 3661 CCCAGTCTCTTTCAGCACCATCAGCACCAGCGCGCCAGGCTCTCCATCTGGCCAGTCCA 3720  
Db 3661 CCCAGTCTCTTTCAGCACCATCAGCACCAGCGCGCCAGGCTCTCCATCTGGCCAGTCCA 3720  
QY 3721 CAGCAGCAGTCAAGCATTTTACACGCGGGGTTGCGCAACTCCCACTCCATGACACCT 3780  
Db 3721 CAGCAGCAGTCAAGCATTTTACACGCGGGGTTGCGCAACTCCCACTCCATGACACCT 3780  
QY 3781 GCCTTCCACAGCAGTCCGACAGAAATAGTTTCCAGCAGCAGCAGCAGTCTCTTTACG 3840  
Db 3781 GCCTTCCACAGCAGTCCGACAGAAATAGTTTCCAGCAGCAGCAGCAGTCTCTTTACG 3840  
QY 3841 ATCCATCTTCTCAGTTTCAGCGCGGTATACCAACCCACCCACATGGCCCACTAGCT 3900  
Db 3841 ATCCATCTTCTCAGTTTCAGCGCGGTATACCAACCCACCCACATGGCCCACTAGCT 3900  
QY 3901 CAGGCTCATGTACAGTCAAGAAATGTTCTCTCTCATCCAACTGCCCCATGCGCCCAATGATG 3960  
Db 3901 CAGGCTCATGTACAGTCAAGAAATGTTCTCTCTCATCCAACTGCCCCATGCGCCCAATGATG 3960  
QY 3961 CTAATGAGCAGCAGCAGCAGCGGGTCCCGAGCGCCCTCGCTCAAAGTGCACTACAG 4020

Db 3961 CTAATGAGCAGCAGCAGCGGGTCCCGAGCGGCTCCCGAGCGCCCTCGCTCAAAGTGCACTACAG 4020  
QY 4021 CCCATTCCAGTCTCGACAACAGCGCATTTCCCTTATATGACGCAACCTTCAGTACAAGCC 4080  
Db 4021 CCCATTCCAGTCTCGACAACAGCGCATTTCCCTTATATGACGCAACCTTCAGTACAAGCC 4080  
QY 4081 CACCACAACAGCAGTGTGTAAGGCTGCCCTGGAGGAACCGGAAGCCCAAAATTCCTCTCTC 4140  
Db 4081 CACCACAACAGCAGTGTGTAAGGCTGCCCTGGAGGAACCGGAAGCCCAAAATTCCTCTCTC 4140  
QY 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGACAGAAAACCTAGAAATTCATTATTTTGTGTTT 4200  
Db 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGACAGAAAACCTAGAAATTCATTATTTTGTGTTT 4200  
QY 4201 TAAATATATATGTTGATTTCTTCTAACTCAATAGGAATGCTTAACAGTTCACTTGCGAG 4260  
Db 4201 TAAATATATATGTTGATTTCTTCTAACTCAATAGGAATGCTTAACAGTTCACTTGCGAG 4260  
QY 4261 TGGAGATACTTGGACCGAGTAGAGGCAATTTAGGAACTTGGGGCTATTCCATATTTCCA 4320  
Db 4261 TGGAGATACTTGGACCGAGTAGAGGCAATTTAGGAACTTGGGGCTATTCCATATTTCCA 4320  
QY 4321 TATCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTTATTT 4380  
Db 4321 TATCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTTATTT 4380  
QY 4381 ATTTTAAATTAACCTTTGAAAGTCATGAACATCAACATCAGCTAGCAAAAGTAACAAGT 4440  
Db 4381 ATTTTAAATTAACCTTTGAAAGTCATGAACATCAACATCAGCTAGCAAAAGTAACAAGT 4440  
QY 4441 GATTCTTGCTGCTATTACTGCTAAAAAAGGCTAAAAAAGGCTAAAAAAGGCTAAAAAAGGCT 4481  
Db 4441 GATTCTTGCTGCTATTACTGCTAAAAAAGGCTAAAAAAGGCTAAAAAAGGCTAAAAAAGGCT 4481

RESULT 2  
AAZ23428 standard; DNA; 4481 BP.  
XX AAZ23428;  
XX 19-JAN-2000 (first entry)  
XX Human SCA2 DNA.  
DE Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;  
KW huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6;  
KW atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;  
KW Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;  
KW dentatorubropallidolysian atrophy; cell proliferation; cell survival;  
KW neoplastic; malignant; autoimmune; fibrotic; ss.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH CDS 163..4101  
FT /\*tag= a  
FT /product= "SCA2"  
XX WO9945944-A1.  
XX 16-SEP-1999.  
XX 11-MAR-1999; 99WO-US005250.  
XX 12-MAR-1998; 98US-00041886.  
XX (BURN-) BURNHAM INST.  
XX Bredesen DE, Rabinzadeh S;  
XX WPI; 1999-561617/47.  
DR P-FSDB; AAY33495.

XX	XX	New proapoptotic dependence peptides, used to develop products for treating, e.g. Alzheimer's disease.
PT	PT	Disclosure; Page 130-135; 199pp; English.
XX	PS	This invention describes novel pure proapoptotic dependence peptides which comprise a sequence of an active dependence domain selected from dependence polypeptides consisting of p75NTR, androgen receptor, DCC, huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2, SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable of inducing cell death and can be used to develop products to mediate or inhibit apoptosis. The methods can be used for reducing the severity of a proapoptotic dependence domain mediated pathological conditions e.g. Huntington's disease, Alzheimer's disease, Kennedy's disease, Spino cerebellar ataxias, dentatorubralpallidoluysian atrophy, Machado-Joseph disease, stroke or head trauma. They can also be used for reducing the severity of a pathological condition mediated by upregulated cell proliferation or cell survival e.g. neoplastic, malignant, autoimmune or fibrotic conditions. This sequence encodes the human SCA2 polypeptide described in the method of the invention
XX	XX	Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
SQL	SQL	Query Match 100.0%; Score 4481; DB 2; Length 4481; Best Local Similarity 100.0%; Pred. No. 0; Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1	ACCCCGAGAAAGCAACCCAGCGCGCCCGCTCCAGTGTCTCCCTCCCGGCCCGGG 60
DB	1	
QY	61	GCACCTCACGTTCTGCTTCGCTCTGACCCCTCCGACTTCCGGTAAGAGTCCCTATCG 120
DB	61	GCACCTCACGTTCTGCTTCGCTCTGACCCCTCCGACTTCCGGTAAGAGTCCCTATCG 120
QY	121	CACCTTCGGTCCACCCGGCGCTCGGGCGCCCGCCTCCGATGCGCTCAGGGCGCGA 180
DB	121	CACCTTCGGTCCACCCGGCGCTCGGGCGCCCGCCTCCGATGCGCTCAGGGCGCGA 180
QY	181	GCTCTCTCGGAGTCCGCGGTGCGCACAGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC 240
DB	181	GCTCTCTCGGAGTCCGCGGTGGCACAGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC 240
QY	241	GGGTGGCGCTCGCTCCAGCGCGCGCGCGAGCGGGCGGGCGCGGTGGCGCGGCC 300
DB	241	GGGTGGCGCTCGCTCCAGCGCGCGCGCGAGCGGGCGGGCGGGTGGCGCGGCC 300
QY	301	CGGGACCGTATCCCTCCGCGCGCCCTCCCGCGCCCGCGCCCGCGCTCCCTCCCGG 360
DB	301	CGGGACCGTATCCCTCCGCGCGCCCTCCCGCGCCCGCGCCCGCGCTCCCTCCCGG 360
QY	361	CAGAGTTCGCTCCCTCCGCGCGCTCTCTGGTCTCGCGCGGGCTCCCGCGCCCTTCGTCGTC 420
DB	361	CAGAGTTCGCTCCCTCCGCGCTCAGACTGTCTTGGTAGCACGCAACGGCGCGCGCGG 420
QY	421	TTTTCGCGCGCGCTCCCGCGGCTCTCTGGTCTCGCGCGGGCTCCCGCGCCCTTCGTCGTC 480
DB	421	TTTTCGCGCGCGCTCCCGCGGCTCTCTGGTCTCGCGCGGGCTCCCGCGCCCTTCGTCGTC 480
QY	481	GTCTTTTCCCCCTTCGCGAGCCGGCGCCCTTCGCGCGCGCGCAACCGCGGCTCCCGG 540
DB	481	GTCTTTTCCCCCTTCGCGAGCCGGCGCCCTTCGCGCGCGCGCAACCGCGGCTCCCGG 540
QY	541	CTCGGCGCCCGTTCGCGCGCGGTTCGCGCGCTCTCTTGGCGCGCGCGGCTCCCGGC 600
DB	541	CTCGGCGCCCGTTCGCGCGCGGTTCGCGCGCTCTCTTGGCGCGCGCGGCTCCCGGC 600
QY	601	TGTCCTCCCGCGCGGTGCGAGCCGGTGTATGGGCCCTTCACATGTCTCGCTGAAGCCCCAG 660
DB	601	TGTCCTCCCGCGCGGTGCGAGCCGGTGTATGGGCCCTTCACATGTCTCGCTGAAGCCCCAG 660
QY	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720
DB	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720



QY 3961 CTAATGACGACACGCCACCGCGGTCCTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020  
DB |||||  
QY 3961 CTAATGACGACACGCCACCGCGGTCCTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020  
DB |||||  
QY 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTCTATATGACGACACCTTTCAGTACAAGCC 4080  
DB |||||  
QY 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTCTATATGACGACACCTTTCAGTACAAGCC 4080  
DB |||||  
QY 4081 CACCAACCAACAGCAGTGTGAAGCTGCGCTGAGGAACCGAAAGCCAAATTCCTCTCTC 4140  
DB |||||  
QY 4081 CACCAACCAACAGCAGTGTGAAGCTGCGCTGAGGAACCGAAAGCCAAATTCCTCTCTC 4140  
DB |||||  
QY 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACAAACTAGAAATTCATTTATTTGTTTT 4200  
DB |||||  
QY 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACAAACTAGAAATTCATTTATTTGTTTT 4200  
DB |||||  
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACATTGCAG 4260  
DB |||||  
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACATTGCAG 4260  
DB |||||  
QY 4261 TGGAAAGATACTTGGACCGAGTAGAGCAATTTAGGAACCTTGGGGCTATTCCATAATTCCA 4320  
DB |||||  
QY 4261 TGGAAAGATACTTGGACCGAGTAGAGCAATTTAGGAACCTTGGGGCTATTCCATAATTCCA 4320  
DB |||||  
QY 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380  
DB |||||  
QY 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380  
DB |||||  
QY 4381 ATTTTATATACCTTTGAAAGTCATGAACATCATGCTAGCAAAAGAGTAAACAGAGT 4440  
DB |||||  
QY 4381 ATTTTATATACCTTTGAAAGTCATGAACATCATGCTAGCAAAAGAGTAAACAGAGT 4440  
DB |||||  
QY 4441 GATTTCTGCTGCTATTACTGCTTAAATTTGCTTAAATTTGCTTAAATTTGCTTAAATTTGCTT 4481  
DB |||||  
QY 4441 GATTTCTGCTGCTATTACTGCTTAAATTTGCTTAAATTTGCTTAAATTTGCTTAAATTTGCTT 4481  
DB |||||

RESULT 3  
ID ADD18753 standard; DNA; 4481 BP.  
XX AC ADD18753;  
XX XT 15-JAN-2004 (first entry)  
XX DE Human disease related protein DNA sequence SeqID184.  
XX KW human; disease state; cytostatic; antiinflammatory; ophthalmological;  
KW antiarteriosclerotic; vulnerary; gene therapy;  
KW hypoxia-regulated condition; tumourigenesis; angiogenesis; apoptosis;  
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing; gene; ds.  
XX OS Homo sapiens.  
XX PN WO2003018621-A2.  
XX PD 06-MAR-2003.  
XX PF 23-AUG-2002; 2002WO-GB003892.  
XX PR 23-AUG-2001; 2001GB-00020558.  
XX PR 05-OCT-2001; 2001GB-00024037.  
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX PI Kingman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
XX WPI; 2003-290046/28.  
XX DR P-PSDB; ADD18752.  
XX DR

XX New substantially purified polypeptide, useful for diagnosing or treating  
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
PT wound healing.  
XX Claim 27; SEQ ID NO 184; 424pp; English.  
XX This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, antiinflammatory,  
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumourigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein encoding DNA sequence of the invention.  
XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 4481; DB 10; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCTCTCACGTGTCCCTCCCGCCCCCGGG 60  
DB |||||  
QY 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCTCTCACGTGTCCCTCCCGCCCCCGGG 60  
DB |||||  
QY 61 GCCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGATTTCCGGTAAAGAGTCCCTATCCG 120  
DB |||||  
QY 61 GCCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGATTTCCGGTAAAGAGTCCCTATCCG 120  
DB |||||  
QY 121 CACCTCCGCTCCACCCCGCGCTCGGCGCGCGCTCCGATGCGGCTCAGGGCGCGCA 180  
DB |||||  
QY 121 CACCTCCGCTCCACCCCGCGCTCGGCGCGCGCTCCGATGCGGCTCAGGGCGCGCA 180  
DB |||||  
QY 181 GCTCTCCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGCCC 240  
DB |||||  
QY 181 GCTCTCCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGCCC 240  
DB |||||  
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCG 300  
DB |||||  
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 300  
DB |||||  
QY 301 CCGGAGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
DB |||||  
QY 301 CCGGAGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
DB |||||  
QY 361 CAGAGTCCGCTCCCTCCGCGCTCAGACTGTTTGTAGCAACGGCAACCGCGCGCGCGCG 420  
DB |||||  
QY 361 CAGAGTCCGCTCCCTCCGCGCTCAGACTGTTTGTAGCAACGGCAACCGCGCGCGCGCG 420  
DB |||||  
QY 421 TTTTCGCGCGCGCTCCCGCGCGCTCTTGTGCTCGGCGGCGCTCCCGCGCGCGCTTCGTCGTC 480  
DB |||||  
QY 421 TTTTCGCGCGCGCTCCCGCGCGCTCTTGTGCTCGGCGGCGCTCCCGCGCGCGCTTCGTCGTC 480  
DB |||||  
QY 481 GTCTTCTCCCTCCAGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB |||||  
QY 481 GTCTTCTCCCTCCAGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB |||||  
QY 541 CTCGCGCGCGCTCGCTCCCGCGCGCTTCGCGGCTCTCTTGGCGCGCGCGCGCGCGCGCG 600  
DB |||||  
QY 541 CTCGCGCGCGCTCGCTCCCGCGCGCTTCGCGGCTCTCTTGGCGCGCGCGCGCGCGCGCG 600  
DB |||||  
QY 601 TGTTCGCGCGCGCGCTCGAGCGCGGTATGGCGCGCTCACCATGTGCTGAAGCGCGCGCG 660  
DB |||||  
QY 601 TGTTCGCGCGCGCGCTCGAGCGCGGTATGGCGCGCTCACCATGTGCTGAAGCGCGCGCG 660  
DB |||||



QY 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720  
Db 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720  
QY 721 CAGCCGCGCGCGCGCGCTGCCAATTCGCGAAGCCCGCGCGCAGCGCGCTTCTAGCGTCG 780  
Db 721 CAGCCGCGCGCGCGCGCTGCCAATTCGCGAAGCCCGCGCGCAGCGCGCTTCTAGCGTCG 780  
QY 781 CCGCGCGCGCGCGCTTCCGCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 840  
Db 781 CCGCGCGCGCGCGCTTCCGCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 840  
QY 841 TCCTCGGTGTCGCGCGCAGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
Db 841 TCCTCGGTGTCGCGCGCAGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
QY 901 AGTAAACAAAGGACTCCCTCAGCTACGATTTCTTTTGATGGAATCTATGCAAAATATGAGG 960  
Db 901 AGTAAACAAAGGACTCCCTCAGCTACGATTTCTTTTGATGGAATCTATGCAAAATATGAGG 960  
QY 961 ATGGTTCAATATCTATCAATCAGTTGTTGGCTCCAAATGTGAAGTCAAGTGAATAATGGA 1020  
Db 961 ATGGTTCAATATCTATCAATCAGTTGTTGGCTCCAAATGTGAAGTCAAGTGAATAATGGA 1020  
QY 1021 GGTATATATGAAGGATTTTAAACTTACAGTCCGAAGTGTGATTTGGTACTTGTATGCC 1080  
Db 1021 GGTATATATGAAGGATTTTAAACTTACAGTCCGAAGTGTGATTTGGTACTTGTATGCC 1080  
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCCGGGCGGAAACGTGAAGAAATATGAGAGGT 1140  
Db 1081 GCACATGAGAAAGTACAGAAATCCAGTTCCGGGCGGAAACGTGAAGAAATATGAGAGGT 1140  
QY 1141 ATTTTGTTCAAATGTTACAGCTTTGTTGGTACAGTTTAAAGATATGAGCTCCAGTTAT 1200  
Db 1141 ATTTTGTTCAAATGTTACAGCTTTGTTGGTACAGTTTAAAGATATGAGCTCCAGTTAT 1200  
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGAAGTGAATGGCGAACAACAA 1260  
Db 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGAAGTGAATGGCGAACAACAA 1260  
QY 1261 GAGAGGACCTGGAGCCCTGGGATCAGTGAACTCACAGCAATGAGGAACCTTGAGGCT 1320  
Db 1261 GAGAGGACCTGGAGCCCTGGGATCAGTGAACTCACAGCAATGAGGAACCTTGAGGCT 1320  
QY 1321 TTGGAATAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGA 1380  
Db 1321 TTGGAATAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGA 1380  
QY 1381 AATTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCTGATATACAGTCCCTTAGAA 1440  
Db 1381 AATTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCTGATATACAGTCCCTTAGAA 1440  
QY 1441 AGAGATAACTCAGAGAATAATTTTAAACGGGAAGCAAGGCAACACAGTTAGCAGAGAA 1500  
Db 1441 AGAGATAACTCAGAGAATAATTTTAAACGGGAAGCAAGGCAACACAGTTAGCAGAGAA 1500  
QY 1501 ATTGAGTCAAGTCCAGTACAAAGCTCCAGTGGCCCTGGAAAATGATGATAGGAGTGAG 1560  
Db 1501 ATTGAGTCAAGTCCAGTACAAAGCTCCAGTGGCCCTGGAAAATGATGATAGGAGTGAG 1560  
QY 1561 GAAGAAAATACAGACGTTCAGAGAAATTCAGTGAACGTGAGGGGCAACAGCATAAAC 1620  
Db 1561 GAAGAAAATACAGACGTTCAGAGAAATTCAGTGAACGTGAGGGGCAACAGCATAAAC 1620  
QY 1621 ACTAGGGAATAATAATATATCTCTCGACAAAGAAATAGAGATCATATCTCGGGGA 1680  
Db 1621 ACTAGGGAATAATAATATATCTCTCGACAAAGAAATAGAGATCATATCTCGGGGA 1680  
QY 1681 AGTGGGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGGCTCCATGCGCATCAAGA 1740  
Db 1681 AGTGGGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGGCTCCATGCGCATCAAGA 1740  
QY 1741 TCCACTTCTCACATTCAGATTTCAACCCGAATTTCTGGTTCAGACCAAAAGAGTAGTTAAT 1800

Db 1741 TCCACTTCTCACATTCAGATTTCAACCCGAATTTCTGGTTCAGACCAAAAGAGTAGTTAAT 1800  
QY 1801 GGAGGTGTTTCCCTGGCCATCGCCTTGCCCATCTCTCTCTCGCCACCTTCTCGCTAC 1860  
Db 1801 GGAGGTGTTTCCCTGGCCATCGCCTTGCCCATCTCTCTCTCGCCACCTTCTCGCTAC 1860  
QY 1861 CAGTCAGGTCCCAACTCTCTTCCACTCTGGGAGCAGCCCTACACGGCGCGCTTCCAGG 1920  
Db 1861 CAGTCAGGTCCCAACTCTCTTCCACTCTGGGAGCAGCCCTACACGGCGCGCTTCCAGG 1920  
QY 1921 CCCCCTCGCGGCGCATCCAGACCCCGCTTCCAGCCCTCTCAGCTCTCTCGAGTCTCT 1980  
Db 1921 CCCCCTCGCGGCGCATCCAGACCCCGCTTCCAGCCCTCTCAGCTCTCTCGAGTCTCT 1980  
QY 1981 GTCTCTACTATGCTTAAACGATGTCTTTCAGAAAGGCGCTCAAGGATGTCCCCAAAGGCC 2040  
Db 1981 GTCTCTACTATGCTTAAACGATGTCTTTCAGAAAGGCGCTCAAGGATGTCCCCAAAGGCC 2040  
QY 2041 CAGCGACATCTCGAAATTCAGAGTTTCTGCTGGAGGGTTCCATATCCAGTGGCCTA 2100  
Db 2041 CAGCGACATCTCGAAATTCAGAGTTTCTGCTGGAGGGTTCCATATCCAGTGGCCTA 2100  
QY 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGAAGCAGT 2160  
Db 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGAAGCAGT 2160  
QY 2161 CCCTCGGGGGAAACGTGTGTCATCAGTGTGAGTGGGTTTCAAGATTTATCCCTTAAACT 2220  
Db 2161 CCCTCGGGGGAAACGTGTGTCATCAGTGTGAGTGGGTTTCAAGATTTATCCCTTAAACT 2220  
QY 2221 CATAGACCCAGGTCTCCAGACAGAAACAGTATTCGAAATACCCCGAGTGGCCAGTCTT 2280  
Db 2221 CATAGACCCAGGTCTCCAGACAGAAACAGTATTCGAAATACCCCGAGTGGCCAGTCTT 2280  
QY 2281 GCTTCTCCCAAGCTGTGTTATTTCCAACTGAAGCTGTGCTCCAGTCTTCCAGCTGCA 2340  
Db 2281 GCTTCTCCCAAGCTGTGTTATTTCCAACTGAAGCTGTGCTCCAGTCTTCCAGCTGCA 2340  
QY 2341 TCTCTACGCTGTAGTCTGTCATCGAAACAGAGCTGTACCCCTTCTAGTGAGGCTAAA 2400  
Db 2341 TCTCTACGCTGTAGTCTGTCATCGAAACAGAGCTGTACCCCTTCTAGTGAGGCTAAA 2400  
QY 2401 GATTCCAGGCTTCAAGTTCAGAGCAGAACTCTCTCGAGGGAATAAGAAAATTTAAA 2460  
Db 2401 GATTCCAGGCTTCAAGTTCAGAGCAGAACTCTCTCGAGGGAATAAGAAAATTTAAA 2460  
QY 2461 CCCAATGAACATCACCTAGTCTTCAAAAGCTGAAAACAAAGGTATATCACAGTTGTT 2520  
Db 2461 CCCAATGAACATCACCTAGTCTTCAAAAGCTGAAAACAAAGGTATATCACAGTTGTT 2520  
QY 2521 TCTCAACATAGAAAACAGATTGATGTTTAAAGAAATTTAAGATGATTTTAGGTTACAG 2580  
Db 2521 TCTCAACATAGAAAACAGATTGATGTTTAAAGAAATTTAAGATGATTTTAGGTTACAG 2580  
QY 2581 CCAAGTTCTACTCTCTGATCTATGATCAACTACTGATCACTAATAACAAAATAGAGAGG 2640  
Db 2581 CCAAGTTCTACTCTCTGATCTATGATCAACTACTAATAACAAAATAGAGAGG 2640  
QY 2641 TCAAGAGATTTGATCAAAAGCAAAATTTGAACCAAGTGTAGGATTTCTTTTCAAGAAAT 2700  
Db 2641 TCAAGAGATTTGATCAAAAGCAAAATTTGAACCAAGTGTAGGATTTCTTTTCAAGAAAT 2700  
QY 2701 AGCAGCAGCACTGACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCTTTCCCTTCA 2760  
Db 2701 AGCAGCAGCACTGACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCTTTCCCTTCA 2760  
QY 2761 ATACTTAGTAAACAGCGGACCAAGAGGGGACCTCAGGTCACTTCCCAAGGGGTTCCAGCT 2820  
Db 2761 ATACTTAGTAAACAGCGGACCAAGAGGGGACCTCAGGTCACTTCCCAAGGGGTTCCAGCT 2820  
QY 2821 TCCAGCCAGCATGTAAACAAAGAAAGACGATGAAGAGAGAAAGACGAGCTGAG 2880

Db 2821 TCAGCCAGCATGTAAACAAGAGAGACGATAGGAAGAGAGAAAGACGCGTGG 2880  
Qy 2881 CAAAGTTAGGAATCAACATTGTAATCCAAATGCAAGAGAGTTCAACCCAGTTCCTTCTCT 2940  
Db 2881 CAAAGTTAGGAATCAACATTGTAATCCAAATGCAAGAGAGTTCAACCCAGTTCCTTCTCT 2940  
Qy 2941 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCATCT 3000  
Db 2941 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCATCT 3000  
Qy 3001 ATGGTGGGTCAATCAACAGCAAACTCCAGTTTATCTCAGCCTGTGTTGTTTGCACCAAT 3060  
Db 3001 ATGGTGGGTCAATCAACAGCAAACTCCAGTTTATCTCAGCCTGTGTTGTTTGCACCAAT 3060  
Qy 3061 ATGATGTATCCAGTCCAGTGCCGCGCGTGCACCTTTATACCCAAATCCTATGAGC 3120  
Db 3061 ATGATGTATCCAGTCCAGTGCCGCGCGTGCACCTTTATACCCAAATCCTATGAGC 3120  
Qy 3121 CCATGCCAGTGAATCAAGCCAGACATATAGAGCAGTACCAATATGCCCAACAGCGG 3180  
Db 3121 CCATGCCAGTGAATCAAGCCAGACATATAGAGCAGTACCAATATGCCCAACAGCGG 3180  
Qy 3181 CAAGCAGCATCATCAGAGTGCATGATGACCCAGCGTTCAGCAGCGGCGCCACCGATT 3240  
Db 3181 CAAGCAGCATCATCAGAGTGCATGATGACCCAGCGTTCAGCAGCGGCGCCACCGATT 3240  
Qy 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCCCTACAGTCTCAGCAGTTCCCA 3300  
Db 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCCCTACAGTCTCAGCAGTTCCCA 3300  
Qy 3301 AATCAGCCCTTGTTCAGATGTCACATATATCAGTCTCAGCATCTCATGCTATAGT 3360  
Db 3301 AATCAGCCCTTGTTCAGATGTCACATATATCAGTCTCAGCATCTCATGCTATAGT 3360  
Qy 3361 CTTGTAATACAGGTTAATGCTAGATGAGGACCAACCAACAGCCAGCTGTTTGA 3420  
Db 3361 CTTGTAATACAGGTTAATGCTAGATGAGGACCAACCAACAGCCAGCTGTTTGA 3420  
Qy 3421 GTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGCAGCATGCGATGATGATG 3480  
Db 3421 GTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGCAGCATGCGATGATGATG 3480  
Qy 3481 CCCAAATACCATACAAAGGAGACAAGCCCTTCTTCTACTTGGCAATTCACGGCG 3540  
Db 3481 CCCAAATACCATACAAAGGAGACAAGCCCTTCTTCTACTTGGCAATTCACGGCG 3540  
Qy 3541 TCCCTTGCTCAGCAGTATGCGACCTTAAGCTTACCTGACCCACATACCTCCACACCT 3600  
Db 3541 TCCCTTGCTCAGCAGTATGCGACCTTAAGCTTACCTGACCCACATACCTCCACACCT 3600  
Qy 3601 CAGCCTTCAGTACCCCACTGACAGCAGCAAGCCAAATGTTGGAAGTCACTCTGCA 3660  
Db 3601 CAGCCTTCAGTACCCCACTGACAGCAGCAAGCCAAATGTTGGAAGTCACTCTGCA 3660  
Qy 3661 CCAGTCTCTGTTACGACCATACAGCAGCGCGCGCAGCTCTCATCTGGCAGTCA 3720  
Db 3661 CCAGTCTCTGTTACGACCATACAGCAGCGCGCGCAGCTCTCATCTGGCAGTCA 3720  
Qy 3721 CAGCAGCAGTACGCAATTTACCAAGCGGGCTTGGCCAACTCCACCTCCATGACACCT 3780  
Db 3721 CAGCAGCAGTACGCAATTTACCAAGCGGGCTTGGCCAACTCCACCTCCATGACACCT 3780  
Qy 3781 GCCTCCAAACAGCAGTCCGACAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTAG 3840  
Db 3781 GCCTCCAAACAGCAGTCCGACAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTAG 3840  
Qy 3841 ATCCATCTCTCAGCTTACGCGGGGTATACCAACCCACCCCAATGCGCCACGTAACCT 3900  
Db 3841 ATCCATCTCTCAGCTTACGCGGGGTATACCAACCCACCCCAATGCGCCACGTAACCT 3900  
Qy 3901 CAGGCTCATGTACAGTACAGAAATGTTCTCTCATCCAACTGCCCATGCGCAATGATG 3960  
Db 3901 CAGGCTCATGTACAGTACAGAAATGTTCTCTCATCCAACTGCCCATGCGCAATGATG 3960

Qy 3961 CTAATGACGACACAGCCACCAGCGGTCCCGAGGCGCGCTCGCTCAAAGTGCACCTACAG 4020  
Db 3961 CTAATGACGACACAGCCACCAGCGGTCCCGAGGCGCGCTCGCTCAAAGTGCACCTACAG 4020  
Qy 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTATATGACGACACCTTTCACTACAAGCC 4080  
Db 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTATATGACGACACCTTTCACTACAAGCC 4080  
Qy 4081 CACCAACCAACAGCAGTGTGTAAGGCTGCCCTGGAGAAACCGAAAGGCCAAATTCCTCTCTC 4140  
Db 4081 CACCAACCAACAGCAGTGTGTAAGGCTGCCCTGGAGAAACCGAAAGGCCAAATTCCTCTCTC 4140  
Qy 4141 CTTCTACTGCTTCTACCAACTGGAAGCAGACAGAAACTAGAAATTTTATTTTGTGTTT 4200  
Db 4141 CTTCTACTGCTTCTACCAACTGGAAGCAGACAGAAACTAGAAATTTTATTTTGTGTTT 4200  
Qy 4201 TAAATATATATGTTGATTCTTGTAAACATCAATAGGAATGCTAACAGTTCACITTCGAG 4260  
Db 4201 TAAATATATATGTTGATTCTTGTAAACATCAATAGGAATGCTAACAGTTCACITTCGAG 4260  
Qy 4261 TGGAAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACTTTGGGGGCTATTTCCTATAATTC 4320  
Db 4261 TGGAAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACTTTGGGGGCTATTTCCTATAATTC 4320  
Qy 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGGAAGTTATTT 4380  
Db 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGGAAGTTATTT 4380  
Qy 4381 ATTTTAAATACCTTGAAGAGTCAATGAAACATCAGCTAGCAAAAGAAAGTAAACAAGAGT 4440  
Db 4381 ATTTTAAATACCTTGAAGAGTCAATGAAACATCAGCTAGCAAAAGAAAGTAAACAAGAGT 4440  
Qy 4441 GATTCCTGCTGCTATTACTGCTTAAATAAAAAAAAAAAAAAAAAAAAA 4481  
Db 4441 GATTCCTGCTGCTATTACTGCTTAAATAAAAAAAAAAAAAAAAAAAAA 4481

## RESULT 4

ADH69466

ID ADH69466 standard; DNA; 4481 BP.

XX AC ADH69466;

XX DT 25-MAR-2004 (first entry)

XX Human spinocerebellar ataxia 2; SCA2; genetic analysis; gene; ds; human.

XX Spinocerebellar ataxia 2; SCA2; genetic analysis; gene; ds; human.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 163..4101

XX /tag= a

XX /product= "Human SCA2 protein"

XX US6623927-B1.

XX 23-SEP-2003.

XX PF 08-NOV-2000; 2000US-00707919.

XX PR 08-NOV-2000; 2000US-00707919.

XX XX (COUL ) COUNCIL SCI &amp; IND RES.

XX PI Brahmachari SK, Choudhry S, Mukerji M, Jain S;

XX DR WPI; 2003-895465/82.

XX P-PSDB; ADH69465.

XX PT New oligonucleotide primers useful in predicting an individual's

PT susceptibility to spinocerebellar ataxia 2 (SCA2) disease, in genetic  
PT analysis of SCA2 gene in a population, and in detecting allelic variants  
PT of the SCA2 gene.

Disclosure; SEQ ID NO 20; 30pp; English.

The invention relates to an oligonucleotide primer. The oligonucleotide primers are useful in molecular diagnosis, in predicting an individual's susceptibility to spinocerebellar ataxia 2 (SCA2) disease, in genetic analysis of SCA2 gene in a population and in detecting allelic variants of the SCA2 gene. The present sequence is human SCA2 gene.

Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

Query Match	Score 4481;	DB 10;	Length 4481;
100.0%			

Seq. ID	Seq. Length	Best Local Similarity	Pred. No. 0;	Mismatches	Indels	Gaps
Matches 4481;	Conservative	0;	0;	0;	0;	0;
100.0%;						

Qy	1	ACCCCGAAGAAAGCAACCCAGCGCGCGCCGCTCTCTACAGTGTCTCTCCCGCGCCCGCGG	60
Db	1	ACCCCGAAGAAAGCAACCCAGCGCGCGCCGCTCTCTACAGTGTCTCTCCCGCGCCCGCGG	60
Qy	61	GCCACCTCAGCTTCGTCTTCGGTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG	120
Db	61	GCCACCTCAGCTTCGTCTTCGGTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG	120
Qy	121	CACCTTCGCTCCACCCCGCGGCTCTCGGCGGCGCCCTCCGATCGCTCAGCGGCGCA	180
Db	121	CACCTTCGCTCCACCCCGCGGCTCTCGGCGGCGCCCTCCGATCGCTCAGCGGCGCA	180
Qy	181	GCTCCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC	240
Db	181	GCTCCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC	240
Qy	241	GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGCGGCGGCGGCGCGCGCGC	300
Db	241	GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGCGGCGGCGGCGCGCGCGC	300
Qy	301	CCGGAGCCGTATCCCTTCGCGCGCCCTCCCGCGCCCGCGCGCGCGCCCTCCCTCCCGG	360
Db	301	CCGGAGCCGTATCCCTTCGCGCGCCCTCCCGCGCCCGCGCGCGCGCCCTCCCTCCCGG	360
Qy	361	CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTTGTAGCAACGGCAACGGCGCGCGCGG	420
Db	361	CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTTGTAGCAACGGCAACGGCGCGCGCGG	420
Qy	421	TTTCGGCCCGGCTCCCGCGGCTCCTTGCTTCGCGCGGCGCTCCCGCGCGCTTCGTCGTC	480
Db	421	TTTCGGCCCGGCTCCCGGCGGCTCCTTGGTCTCGGCGGCGCTCCCGCGCGCTTCGTCGTC	480
Qy	481	GTCCTTTCCTCCCTTCGCGCAGCGCGGCGCCCTTCGCGCGCGGCCAACCGCGCGCTCCCGG	540
Db	481	GTCCTTTCCTCCCTTCGCGCAGCGCGGCGCCCTTCGCGCGCGGCCAACCGCGCGCTCCCGG	540
Qy	541	CTCGCGCGCGTGGCTCCCGCGGTTCCGCGCTCTCTTGGCGGCGCGCGCTCCCGG	600
Db	541	CTCGCGCGCGTGGCTCCCGCGGTTCCGCGCTCTCTTGGCGGCGCGCGCTCCCGG	600
Qy	601	TGTCCTCCCGCGCGGTGCGAGCCGGTGTATGGGCGCCCTCACCATGTGCTGAAGCGCCAG	660
Db	601	TGTCCCGCGCGGTGCGAGCCGGTGTATGGGCGCCCTCACCATGTGCTGAAGCGCCAG	660
Qy	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
Db	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
Qy	721	CAGCGCGCGCGCGGTGCCAATGTCCGCAAGCCCGCGCGCGCGCTTCTAGCGTCG	780
Db	721	CAGCGCGCGCGCGGTGCCAATGTCCGCAAGCCCGCGCGCGCGCTTCTAGCGTCG	780
Qy	781	CCCGCGCGCGCGCTTCGCGCGTCTCTCGTCTCGGCTCTCTCGTCTCGGCGCACGCTCC	840
Db	781	CCCGCGCGCGCGCTTCGCGCGTCTCTCGTCTCGGCTCTCTCGTCTCGGCGCACGCTCC	840

QY 1921 CCGCCTCGGGCCATCCAGACCCCGCTCTCACCCCTCTGCTCATGTTCTCCAGCTCT 1980  
DB |||||  
QY 1921 CCGCCTCGGGCCATCCAGACCCCGCTCTCACCCCTCTGCTCATGTTCTCCAGCTCT 1980  
DB |||||  
QY 1981 GTCTCTACTATGCTTAAAGCCATGCTCTCAGAAAGGCGCTTCAAGGATGTCCTCCAAAGGCC 2040  
DB |||||  
QY 1981 GTCTCTACTATGCTTAAAGCCATGCTCTCAGAAAGGCGCTTCAAGGATGTCCTCCAAAGGCC 2040  
DB |||||  
QY 2041 CAGCGACATCTCGAAATCAGAGATTTCTGTGGAGGGGTTCCATATCCAGTGGCCTTA 2100  
DB |||||  
QY 2041 CAGCGACATCTCGAAATCAGAGATTTCTGTGGAGGGGTTCCATATCCAGTGGCCTTA 2100  
DB |||||  
QY 2101 GAATTTGTATCCAGACCCACCCAGTGAAGCAGTACTCTCCAGTGAAGGACCACT 2160  
DB |||||  
QY 2101 GAATTTGTATCCAGACCCACCCAGTGAAGCAGTACTCTCCAGTGAAGGACCACT 2160  
DB |||||  
QY 2161 CCTCGGGGGGAAAGTGGTCTCATGAGTGGTTCAGTGGGGTTCCAGATTAATCCCTTAAACT 2220  
DB |||||  
QY 2161 CCTCGGGGGGAAAGTGGTCTCATGAGTGGTTCAGTGGGGTTCCAGATTAATCCCTTAAACT 2220  
DB |||||  
QY 2221 CATAGACCCAGGTCTCCAGACAGAACAGTATTTGGAAATACCCCAAGTGGGCGAGTTCTT 2280  
DB |||||  
QY 2221 CATAGACCCAGGTCTCCAGACAGAACAGTATTTGGAAATACCCCAAGTGGGCGAGTTCTT 2280  
DB |||||  
QY 2281 GCTTCTCCCAAGCTGGTATTTCAACTGAAGCTGTTGCCATGCTATTCCAGCTGCA 2340  
DB |||||  
QY 2281 GCTTCTCCCAAGCTGGTATTTCAACTGAAGCTGTTGCCATGCTATTCCAGCTGCA 2340  
DB |||||  
QY 2341 TCTCTCTACGCTCTAGTCTGATCGAAGCAGAGCTGTTACCCCTCTTAGTGAGGCTAAA 2400  
DB |||||  
QY 2341 TCTCTCTACGCTCTAGTCTGATCGAAGCAGAGCTGTTACCCCTCTTAGTGAGGCTAAA 2400  
DB |||||  
QY 2401 GAATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCGAGGGAATAAGAAATATTTAAA 2460  
DB |||||  
QY 2401 GAATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCGAGGGAATAAGAAATATTTAAA 2460  
DB |||||  
QY 2461 CCCAATGAACATCAGCTAGTCTTCAAGAGCTGAAACAAAGGTATATCACAGTTGTT 2520  
DB |||||  
QY 2461 CCCAATGAACATCAGCTAGTCTTCAAGAGCTGAAACAAAGGTATATCACAGTTGTT 2520  
DB |||||  
QY 2521 TCTGAAATGAAGAAACAGATGATGATTTAAAGAAATTTAAGAAATGATTTAGTTACAG 2580  
DB |||||  
QY 2521 TCTGAAATGAAGAAACAGATGATGATTTAAAGAAATTTAAGAAATGATTTAGTTACAG 2580  
DB |||||  
QY 2581 CCAAGTTCTACTTCTGAATCTATGATCAACTATAACAAATAAGAGGGAGAGAAA 2640  
DB |||||  
QY 2581 CCAAGTTCTACTTCTGAATCTATGATCAACTATAACAAATAAGAGGGAGAGAAA 2640  
DB |||||  
QY 2641 TCAAGAGATTTGATCAAGAACAAATTTGAACCAAGTCTGATGATTTCTTCAATGAAAT 2700  
DB |||||  
QY 2641 TCAAGAGATTTGATCAAGAACAAATTTGAACCAAGTCTGATGATTTCTTCAATGAAAT 2700  
DB |||||  
QY 2701 AGCAGCAGCACTGTACCGTGCAGCAGCAAGCCGAATAGCCCGAGCATTTCCCTTCA 2760  
DB |||||  
QY 2701 AGCAGCAGCACTGTACCGTGCAGCAGCAAGCCGAATAGCCCGAGCATTTCCCTTCA 2760  
DB |||||  
QY 2761 ATACTTAGTAAACGGAGCAAGAGGGGACCTGAGTCACTTTCCAAAGGGGTTGAGCT 2820  
DB |||||  
QY 2761 ATACTTAGTAAACGGAGCAAGAGGGGACCTGAGTCACTTTCCAAAGGGGTTGAGCT 2820  
DB |||||  
QY 2821 TCCAGGCCAGCATGTAAACAGAGAAAGCAGATTAAGAGAGAGAAAGACGCGAGCTGAG 2880  
DB |||||  
QY 2821 TCCAGGCCAGCATGTAAACAGAGAAAGCAGATTAAGAGAGAGAAAGACGCGAGCTGAG 2880  
DB |||||  
QY 2881 CRAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTCTCTCTCT 2940  
DB |||||  
QY 2881 CRAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTCTCTCTCT 2940  
DB |||||  
QY 2941 CAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAAGCACAACCTAGCCCATCT 3000  
DB |||||  
QY 2941 CAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAAGCACAACCTAGCCCATCT 3000  
DB |||||  
QY 3001 ATGGTGGGTCAATCAAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTGGCACCAAT 3060

DB |||||  
QY 3001 ATGGTGGGTCAATCAAGCCAACTCCAGTTTATATCTCAGCCTGTTGTTTGGCACCAAT 3060  
DB |||||  
QY 3061 ATGATGTATCCAGTCCCAGTGAGCCAGCGGTGAAACCTTTTATACCAATACCTATGAGC 3120  
DB |||||  
QY 3061 ATGATGTATCCAGTCCCAGTGAGCCAGCGGTGAAACCTTTTATACCAATACCTATGAGC 3120  
DB |||||  
QY 3121 CCCATGCCAGTGAATCAAGCCAAAGCATATAGAGCAGTACCAAAATATGCCCCAACAGCGG 3180  
DB |||||  
QY 3121 CCCATGCCAGTGAATCAAGCCAAAGCATATAGAGCAGTACCAAAATATGCCCCAACAGCGG 3180  
DB |||||  
QY 3181 CAAAGCCAGCATCATCAGAGTCCATGATGACCCAGCGTTCAGCAGCGGCCACCGATTT 3240  
DB |||||  
QY 3181 CAAAGCCAGCATCATCAGAGTCCATGATGACCCAGCGTTCAGCAGCGGCCACCGATTT 3240  
DB |||||  
QY 3241 GCAGCCACCCACAGCTTACTCCACGCAATATGTTGCCCTACAGTCTCAGCAGTTCCTCA 3300  
DB |||||  
QY 3241 GCAGCCACCCACAGCTTACTCCACGCAATATGTTGCCCTACAGTCTCAGCAGTTCCTCA 3300  
DB |||||  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCAATATCAGTCTCAGCATCCTCATCTATAGT 3360  
DB |||||  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCAATATCAGTCTCAGCATCCTCATCTATAGT 3360  
DB |||||  
QY 3361 CCTGTAAATACAGGGTAATGCTAGAATGATGGCACCAACACACAGCCCGAGCTTGT 3420  
DB |||||  
QY 3361 CCTGTAAATACAGGGTAATGCTAGAATGATGGCACCAACACACAGCCCGAGCTTGT 3420  
DB |||||  
QY 3421 GTATCTTCTTACAGCACTCAGTACGGGGCTATGAGCAGACGATGCGATGATGATGT 3480  
DB |||||  
QY 3421 GTATCTTCTTACAGCACTCAGTACGGGGCTATGAGCAGACGATGCGATGATGATGT 3480  
DB |||||  
QY 3481 CCCAAATTTACCATCAACAAAGGAGCAAGCCCTTCTTTTACTTTGCGCATTTCCAGGGC 3540  
DB |||||  
QY 3481 CCCAAATTTACCATCAACAAAGGAGCAAGCCCTTCTTTTACTTTGCGCATTTCCAGGGC 3540  
DB |||||  
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCTAAGCTTACCTCAGCCACACATACTCCACACCT 3600  
DB |||||  
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCTAAGCTTACCTCAGCCACACATACTCCACACCT 3600  
DB |||||  
QY 3601 CAGCCTTACAGCTACCCCTCAGCAGCAGCAAGCAACATGTTGGTGAAGTCACTCTGCA 3660  
DB |||||  
QY 3601 CAGCCTTACAGCTACCCCTCAGCAGCAGCAAGCAACATGTTGGTGAAGTCACTCTGCA 3660  
DB |||||  
QY 3661 CCGAGTCTGTTTACAGCACCATCAGCAGCAGCGGGCTCCTCCTCTGCGCATTTCCAGGGC 3720  
DB |||||  
QY 3661 CCGAGTCTGTTTACAGCACCATCAGCAGCAGCGGGCTCCTCCTCTGCGCATTTCCAGGGC 3720  
DB |||||  
QY 3721 CAGCAGCAGTACAGCCTTTTACCAAGCGGGCTTGGCCAACTCCACCTCCATGACACCT 3780  
DB |||||  
QY 3721 CAGCAGCAGTACAGCCTTTTACCAAGCGGGCTTGGCCAACTCCACCTCCATGACACCT 3780  
DB |||||  
QY 3781 GCTCTCAACACAGCAGTCCGACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
DB |||||  
QY 3781 GCTCTCAACACAGCAGTCCGACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
DB |||||  
QY 3841 ATCCATCTTCTCAGCTTACAGCGGGCTATACCAACCCACCCACATGCGCCACAGTACCT 3900  
DB |||||  
QY 3841 ATCCATCTTCTCAGCTTACAGCGGGCTATACCAACCCACCCACATGCGCCACAGTACCT 3900  
DB |||||  
QY 3901 CAGGCTCATGATCAGTCAGGAATGTTCTTCTCATCCAACTCCCATCGCCCAATGATG 3960  
DB |||||  
QY 3901 CAGGCTCATGATCAGTCAGGAATGTTCTTCTCATCCAACTCGCCCAATGATG 3960  
DB |||||  
QY 3961 CTAAATGACGACACAGCAGCAGCGGGTCCCGAGCGCGCTCTCGCTCAAAAGTGCATACAG 4020  
DB |||||  
QY 3961 CTAAATGACGACACAGCAGCAGCGGGTCCCGAGCGCGCTCTCGCTCAAAAGTGCATACAG 4020  
DB |||||  
QY 4021 CCAATTTCCAGTCTCGACAAAGCGCATTTTCCCTTATATGAGCAGCCCTTTCAGTACAAGCC 4080  
DB |||||  
QY 4021 CCAATTTCCAGTCTCGACAAAGCGCATTTTCCCTTATATGAGCAGCAGCCCTTTCAGTACAAGCC 4080  
DB |||||  
QY 4081 CACCAACACAGCAGTGTAAAGCTGCTTGGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140

Db 4081 CACCACCAACAGCAGTGTGTAAGGCTGCCTGGAGAACCGAAAGGCCAAATTCCTCTCTC 4140  
Qy 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACGAAACTAGAAATTTCAATTTATTTGTTTT 4200  
Db 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACGAAACTAGAAATTTCAATTTATTTGTTTT 4200  
Qy 4201 TAAAAATATATGTTGATTCTTGTAAACATCCAATAGGAATGCTTAACAGTTTCACTTGCAG 4260  
Db 4201 TAAAAATATATGTTGATTCTTGTAAACATCCAATAGGAATGCTTAACAGTTTCACTTGCAG 4260  
Qy 4261 TGGAGATACTTGGACCGAGTAGAGGCATTTAGGAACTTGGGGGCTATTCCATAATTTCCA 4320  
Db 4261 TGGAGATACTTGGACCGAGTAGAGGCATTTAGGAACTTGGGGGCTATTCCATAATTTCCA 4320  
Qy 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGC CGGAAA CTGGAA GTTATTT 4380  
Db 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGC CGGAAA CTGGAA GTTATTT 4380  
Qy 4381 ATTTTTTAATACCCCTTGAAAGTCAATGAACATCAGCTAGCMAAAGAAAGTAACAAGAGT 4440  
Db 4381 ATTTTTTAATACCCCTTGAAAGTCAATGAACATCAGCTAGCMAAAGAAAGTAACAAGAGT 4440  
Qy 4441 GATTCCTGCTGCTATTACTGTATAAAAAAAAAAAAAAAAAAAAA 4481  
Db 4441 GATTCCTGCTGCTATTACTGTATAAAAAAAAAAAAAAAAAAAAA 4481

RESULT 5

ABV75291  
ID ABV75291 standard; DNA; 4481 BP.  
XX  
AC ABV75291;  
XX  
DT 07-MAR-2003 (first entry)  
XX  
DE Human SCA2 polypeptide encoding DNA.  
XX  
KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;  
KW memory impairment; mutant; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 163..4101  
FT /\*tag= a  
FT /\*product= "SCA2"  
XX  
PN WO200289567-A2.  
XX  
PD 14-NOV-2002.  
XX  
PF 07-MAY-2002; 2002WO-US014698.  
XX  
PR 07-MAY-2001; 2001US-0289231P.  
XX  
PA (CEDA-) CEDARS SINAI MEDICAL CENT.  
XX  
PI Pulist SM;  
XX  
WPI; 2003-103467/09.  
DR P-PSDB; ABB82698.  
XX  
PT New mutant mouse, useful as a disease model for identifying agents for  
PT treating obesity or memory impairment, comprises a disrupted  
PT spinocerebellar ataxia 2 gene.  
XX  
PS Disclosure; Page 65-72; 76pp; English.  
XX

The invention relates to a mutant non-human mammal having a disrupted spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a disease model for identifying agents or compounds that may treat obesity or memory impairment. The stem cell is used for generating the mutant animal. The present sequence represents the nucleotide sequence (GenBank

CC Accession No. U70323) encoding the human SCA2 polypeptide  
XX  
SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;  
Query Match 100.0%; Score 4481; DB 10; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACCCCCAGAGAAAGCAACCCAGCGCGCCGCTCTCTCACTGCGGTAAAGAGTCCCTATCCG 120  
Db 1 ACCCCCAGAGAAAGCAACCCAGCGCGCGCCGCTCTCTCACTGCGGTAAAGAGTCCCTATCCG 120  
Qy 61 GCCACCTCACGTTCTGCTTCCGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120  
Db 61 GCCACCTCACGTTCTGCTTCCGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120  
Qy 121 CACCTCCGCTCCCAACCCGCGGCCCTCGCGCGCCGCTCCGATGCGCTCAGCGCGCCCA 180  
Db 121 CACCTCCGCTCCCAACCCGCGGCCCTCGCGCGCCGCTCCGATGCGCTCAGCGCGCCCA 180  
Qy 181 GCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC 240  
Db 181 GCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC 240  
Qy 241 GGGTGGCGCTCGCTCCAGCGGCCGCGCGGAGCGGGCGGGCGGGTGGCGCGGCC 300  
Db 241 GGGTGGCGCTCGCTCCAGCGGCCGCGCGGAGCGGGCGGGCGGGTGGCGCGGCC 300  
Qy 301 CCGGAGCGGTATCCCTCCGCGGCCCTCCCGCGGCCGCGGCCGCGGCCCTCCCTCCCGG 360  
Db 301 CCGGAGCGGTATCCCTCCGCGGCCCTCCCGCGGCCGCGGCCGCGGCCCTCCCTCCCGG 360  
Qy 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAGCGCGCGCGCG 420  
Db 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAGCGCGCGCGCG 420  
Qy 421 TTTTGGGCGCGCTCGGGCGGCTCTTGGTCTCGGGGGGCTCCCGGCCCTTCTGTCGTG 480  
Db 421 TTTTGGGCGCGCTCGGGCGGCTCTTGGTCTCGGGGGGCTCCCGGCCCTTCTGTCGTG 480  
Qy 481 GTCTTCTCCCTCCGCGCGGCCCTCCCGCGGCCGCGCAACCGCGCGCTCCCGG 540  
Db 481 GTCTTCTCCCTCCGCGCGGCCCTCCCGCGGCCGCGCAACCGCGCGCTCCCGG 540  
Qy 541 CTGCGCGCGGTGCGTCCCGCGCGCTTCCGCGCGCTCTCTTGGCGCGCGCGGCTCCGCG 600  
Db 541 CTGCGCGCGGTGCGTCCCGCGCGCTTCCGCGCGCTCTCTTGGCGCGCGCGGCTCCGCG 600  
Qy 601 TGTCCCGCGCGGTGCGAGCCGCTGTATGGGGCCCTCACCATGTGCTGAAGCCCGCAG 660  
Db 601 TGTCCCGCGCGGTGCGAGCCGCTGTATGGGGCCCTCACCATGTGCTGAAGCCCGCAG 660  
Qy 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAG 720  
Db 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAG 720  
Qy 721 CAGCGCGCGCGGCTGCGCAATGTCCGAAAGCCCGCGCGCGCGCGCTTCTAGCGGTG 780  
Db 721 CAGCGCGCGCGGCTGCGCAATGTCCGAAAGCCCGCGCGCGCGCGCTTCTAGCGGTG 780  
Qy 781 CCGCGCGCGCGCTTCCGCGCTCGTCTCGGTCTCTGCTCTGCGCCACCGGCTCCC 840  
Db 781 CCGCGCGCGCGCTTCCGCGCTCGTCTCGGTCTCTGCTCTGCGCCACCGGCTCCC 840  
Qy 841 TCCTCGGTGTCGCGCGACCTCCGCGCGGGAGGCCCGGCCCTGGGAGAGGTGGAAC 900  
Db 841 TCCTCGGTGTCGCGCGACCTCCGCGCGGGAGGCCCGGCCCTGGGAGAGGTGGAAC 900  
Qy 901 AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGTATGGAATCTATGCAAAATATAGG 960  
Db 901 AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGTATGGAATCTATGCAAAATATAGG 960  
Qy 961 ATGGTTCAATATACTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTCAAAATATGA 1020

||||| 961 ATGGTTCATATACCTACATCAGTTGTGGCTCCAAATGTGAAGTACAAAGTGAATAATGGA 1020  
QY  
1021 GGTATATGAAGGAGTGTAACTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080  
Db  
1021 GGTATATGAAGGAGTGTAACTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080  
QY  
1081 GCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGTGAAGAAATATGAGAGT 1140  
Db  
1081 GCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGTGAAGAAATATGAGAGT 1140  
QY  
1141 ATTTTGTTCAAATGTTCCAGCTTTCTGTGGTACAGTTTAAAGATATGACTCCAGTTAT 1200  
Db  
1141 ATTTTGTTCAAATGTTCCAGCTTTCTGTGGTACAGTTTAAAGATATGACTCCAGTTAT 1200  
QY  
1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGCGAAACACAAA 1260  
Db  
1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGCGAAACACAAA 1260  
QY  
1261 GAGAAAGCCTGGAGCCTGGAGTGGATGAGTGAATCTCAGCCAAATGAGGAACTTGAGGCT 1320  
Db  
1261 GAGAAAGCCTGGAGCCTGGAGTGGATGAGTGAATCTCAGCCAAATGAGGAACTTGAGGCT 1320  
QY  
1321 TTGGAAATGAGCTATCTAAATGGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380  
Db  
1321 TTGGAAATGAGCTATCTAAATGGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380  
QY  
1381 AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCTATACAGTGCCTTAGAA 1440  
Db  
1381 AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCTATACAGTGCCTTAGAA 1440  
QY  
1441 AGAGATAACTCAGAGAAATTTTAAACCGGGAAGCAAGGGCAAAACAGTTAGCAGAA 1500  
Db  
1441 AGAGATAACTCAGAGAAATTTTAAACCGGGAAGCAAGGGCAAAACAGTTAGCAGAA 1500  
QY  
1501 ATTGAGTCAAGTCCAGTACAAAGCTCAGTGGCCCTGGAAATGATGATAGGAGTGA 1560  
Db  
1501 ATTGAGTCAAGTCCAGTACAAAGCTCAGTGGCCCTGGAAATGATGATAGGAGTGA 1560  
QY  
1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC 1620  
Db  
1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC 1620  
QY  
1621 ACTAGGAAAAATAAATATATTCCTCTGCAAAAGAAATAGAGAGTATATCTCTGGGA 1680  
Db  
1621 ACTAGGAAAAATAAATATATTCCTCTGCAAAAGAAATAGAGAGTATATCTCTGGGA 1680  
QY  
1681 AGTGGGAGACAGAAATTCACCGGTATGGGCGCAGCCTGGATCGGGCTCCATGCCATCAGA 1740  
Db  
1681 AGTGGGAGACAGAAATTCACCGGTATGGGCGCAGCCTGGATCGGGCTCCATGCCATCAGA 1740  
QY  
1741 TCCACTTCTCACACTTCAGATTCACACCGAAATTCCTGGTTTCAGACCAAGAGTAGTTAT 1800  
Db  
1741 TCCACTTCTCACACTTCAGATTCACACCGAAATTCCTGGTTTCAGACCAAGAGTAGTTAT 1800  
QY  
1801 GGAGTGTTCCTGGCCATCGCTTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
Db  
1801 GGAGTGTTCCTGGCCATCGCTTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
QY  
1861 CAGTCAAGTCCCAACTCTCTTCCACCTCGGGGAGCCACCTTACAGCGGGCCCTCCAGG 1920  
Db  
1861 CAGTCAAGTCCCAACTCTCTTCCACCTCGGGGAGCCACCTTACAGCGGGCCCTCCAGG 1920  
QY  
1921 CCCCCCTCGGGGCATCCAGACCCCGTCTCACCTCTGTCTCATCTGGTTCTCAGCTCT 1980  
Db  
1921 CCCCCCTCGGGGCATCCAGACCCCGTCTCACCTCTGTCTCATCTGGTTCTCAGCTCT 1980  
QY  
1981 GTCTCTACTATGCTTAAACGCGATGCTTTCAGAGGGCCTCCAGAGTGTCCCAAGGCC 2040  
Db  
1981 GTCTCTACTATGCTTAAACGCGATGCTTTCAGAGGGCCTCCAGAGTGTCCCAAGGCC 2040  
QY  
2041 CAGCGACATCTCGAAATCACAGAGTTTCTGTGGGAGGGGTTCCATATCCAGTGGCCTA 2100  
|||||

Db  
2041 CAGCGACATCTCGAAATCACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100  
QY  
2101 GAATTTGTATCCCAACACCCACCAGGTGAAGAGCTACTCTCTCAGTAGCAAGACCACT 2160  
Db  
2101 GAATTTGTATCCCAACACCCACCAGGTGAAGAGCTACTCTCTCAGTAGCAAGACCACT 2160  
QY  
2161 CCCTCGGGGGAAACGTGTGTCATCAGTGTGTGAGTGGGTTTCCAAAGATTATCCCTCAAACT 2220  
Db  
2161 CCCTCGGGGGAAACGTGTGTCATCAGTGTGTGAGTGGGTTTCCAAAGATTATCCCTCAAACT 2220  
QY  
2221 CATAGACCCAGGTCTCCAGACAGAAACAGTATTTGAAATACCCCCAGTGGGCGAGTTCT 2280  
Db  
2221 CATAGACCCAGGTCTCCAGACAGAAACAGTATTTGAAATACCCCCAGTGGGCGAGTTCT 2280  
QY  
2281 GCTTCTCCCAAGCTGTGATTTATTTCCAACTGAAGCTGTTCGCAATGCTTATTCACGTGCA 2340  
Db  
2281 GCTTCTCCCAAGCTGTGATTTATTTCCAACTGAAGCTGTTCGCAATGCTTATTCACGTGCA 2340  
QY  
2341 TCTCTACGCTGTAGTCTGCAATCGAAGAGCTGTACCCCTTCTAGTGAGGCTAAA 2400  
Db  
2341 TCTCTACGCTGTAGTCTGCAATCGAAGAGCTGTACCCCTTCTAGTGAGGCTAAA 2400  
QY  
2401 GATTCCAGGCTTCAAGATCAGAGGAGCAAACTCTCTGAGGGGAAATAAAGAAATATTTAAA 2460  
Db  
2401 GATTCCAGGCTTCAAGATCAGAGGAGCAAACTCTCTGAGGGGAAATAAAGAAATATTTAAA 2460  
QY  
2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAGTTGTT 2520  
Db  
2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAGTTGTT 2520  
QY  
2521 TCTGAACATAGAAACAGATGATGATTTAAAGAAATTTAAAGAAATTTAAAGATTTAGTTACAG 2580  
Db  
2521 TCTGAACATAGAAACAGATGATGATTTAAAGAAATTTAAAGAAATTTAAAGATTTAGTTACAG 2580  
QY  
2581 CCAAGTCTTACTTCTGAATCTATGGATCAATCTACTAAACAAATAAGAGAGGAGAA 2640  
Db  
2581 CCAAGTCTTACTTCTGAATCTATGGATCAATCTACTAAACAAATAAGAGAGGAGAA 2640  
QY  
2641 TCAAGAGATTTGATCAAAAGACAAATTTGAACTCAAGTGTCAAGGATTTCTTTCAATGAAAT 2700  
Db  
2641 TCAAGAGATTTGATCAAAAGACAAATTTGAACTCAAGTGTCAAGGATTTCTTTCAATGAAAT 2700  
QY  
2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAGCGAATAGCCCGAGCATTTCCCTCTCA 2760  
Db  
2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAGCGAATAGCCCGAGCATTTCCCTCTCA 2760  
QY  
2761 ATACTTAGTAAACACGGAGCACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA 2820  
Db  
2761 ATACTTAGTAAACACGGAGCACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA 2820  
QY  
2821 TCCAGCCCGAGCATGTAAACAGAGAAACGATAGAGAGAGAAAGAGAGAGAGAGAGAGAG 2880  
Db  
2821 TCCAGCCCGAGCATGTAAACAGAGAAACGATAGAGAGAGAAAGAGAGAGAGAGAGAGAG 2880  
QY  
2881 CAAAGTTAGGAAATCAACATTGAAATCCCAATGCAAGAGGAGTTCAACCCAGTTCTCT 2940  
Db  
2881 CAAAGTTAGGAAATCAACATTGAAATCCCAATGCAAGAGGAGTTCAACCCAGTTCTCT 2940  
QY  
2941 CAGCCAAAGCCTTCTACTACCCCAACTTCCACCTCGGCTCAAGCAGCAAACTAGCCCCATCT 3000  
Db  
2941 CAGCCAAAGCCTTCTACTACCCCAACTTCCACCTCGGCTCAAGCAGCAAACTAGCCCCATCT 3000  
QY  
3001 ATGGTGGGTCACTCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTGGACCAAT 3060  
Db  
3001 ATGGTGGGTCACTCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTGGACCAAT 3060  
QY  
3061 ATGATGTATCCAGTCCAGTGAGCCGAGCGTGCAACCTTTATACCCAATACCTATGAGC 3120  
Db  
3061 ATGATGTATCCAGTCCAGTGAGCCGAGCGTGCAACCTTTATACCCAATACCTATGAGC 3120  
QY  
3121 CCCATGCCAGTCAATCAAGCCCAAGCATATAGAGCAGTACCAAAATATGCCCCAACAGCGG 3180  
Db  
3121 CCCATGCCAGTCAATCAAGCCCAAGCATATAGAGCAGTACCAAAATATGCCCCAACAGCGG 3180

QY 3181 CAAGACGAGCATCATCAGAGTGCATGATGACCCAGCGTCAGCAGCGGGCCACCGGATT 3240  
DB |||||  
QY 3181 CAAGACGAGCATCATCAGAGTGCATGATGACCCAGCGTCAGCAGCGGGCCACCGGATT 3240  
DB |||||  
QY 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTTACAGTCCTCAGCAGTTCCCA 3300  
DB |||||  
QY 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTTACAGTCCTCAGCAGTTCCCA 3300  
DB |||||  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCATGTCTATAGT 3360  
DB |||||  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCATGTCTATAGT 3360  
DB |||||  
QY 3361 CCTGTAATACAGGGTAATGCTAGATATGATGACCAACACAGCGCCAGCTGGTTTA 3420  
DB |||||  
QY 3361 CCTGTAATACAGGGTAATGCTAGATATGATGACCAACACAGCGCCAGCTGGTTTA 3420  
DB |||||  
QY 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATGTATGCATGT 3480  
DB |||||  
QY 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATGTATGCATGT 3480  
DB |||||  
QY 3481 CCCAAATTACCATCAACAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540  
DB |||||  
QY 3481 CCCAAATTACCATCAACAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540  
DB |||||  
QY 3541 TCCCTTGTCTCAGCATGTGCGACACCTTAACGCTACCTTGCAACCATATCTCCACACCT 3600  
DB |||||  
QY 3541 TCCCTTGTCTCAGCATGTGCGACACCTTAACGCTACCTTGCAACCATATCTCCACACCT 3600  
DB |||||  
QY 3601 CAGCCTTCAGTACCCCACTGGACAGCAGCAAGCCAAACATGTTGGAAGTCATCTCTCA 3660  
DB |||||  
QY 3601 CAGCCTTCAGTACCCCACTGGACAGCAGCAAGCCAAACATGTTGGAAGTCATCTCTCA 3660  
DB |||||  
QY 3661 CCCAGTCTCTTCAGCACCATCAGCACCGCCGCCAGGCTCTCATCTGCGCAGTCCA 3720  
DB |||||  
QY 3661 CCCAGTCTCTTCAGCACCATCAGCACCGCCGCCAGGCTCTCATCTGCGCAGTCCA 3720  
DB |||||  
QY 3721 CAGCAGCAGTACGACCATTTACACCGGGGGTTGCGCAACTCCACCTCCATGACACCT 3780  
DB |||||  
QY 3721 CAGCAGCAGTACGACCATTTACACCGGGGGTTGCGCAACTCCACCTCCATGACACCT 3780  
DB |||||  
QY 3781 GCCTCCACACGCGTCCGACAGATAGTTTCCGAGCAGCACAACAGACTGTCTTTACG 3840  
DB |||||  
QY 3781 GCCTCCACACGCGTCCGACAGATAGTTTCCGAGCAGCACAACAGACTGTCTTTACG 3840  
DB |||||  
QY 3841 ATCCATCTCTTCAGCTTCAGCGCGGTATACCAACCCACCCACATGCGCCACGTACT 3900  
DB |||||  
QY 3841 ATCCATCTCTTCAGCTTCAGCGCGGTATACCAACCCACCCACATGCGCCACGTACT 3900  
DB |||||  
QY 3901 CAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCCAACTGCGCCCAATGATG 3960  
DB |||||  
QY 3901 CAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCCAACTGCGCCCAATGATG 3960  
DB |||||  
QY 3961 CTATGAGCAGACACGCCACCGCGGTCCAGGCGCGCTCGCTCAAAGTGCATACAG 4020  
DB |||||  
QY 3961 CTATGAGCAGACACGCCACCGCGGTCCAGGCGCGCTCGCTCAAAGTGCATACAG 4020  
DB |||||  
QY 4021 CCCATTCCAGTCTCAGCAACAGCGATTTCCCTTATATGACGACCTTTAGTACAAGCC 4080  
DB |||||  
QY 4021 CCCATTCCAGTCTCAGCAACAGCGATTTCCCTTATATGACGACCTTTAGTACAAGCC 4080  
DB |||||  
QY 4081 CACCAACCAACAGCAGTGTAAAGTGCCTGGAGAAACCGAAAGGCCAAATTCCTCTC 4140  
DB |||||  
QY 4081 CACCAACCAACAGCAGTGTAAAGTGCCTGGAGAAACCGAAAGGCCAAATTCCTCTC 4140  
DB |||||  
QY 4141 CCTTCTACTGCTTCTACCACTGGAGCAGCAAAACTAGATTTTCAATTTTGTGTTT 4200  
DB |||||  
QY 4141 CCTTCTACTGCTTCTACCACTGGAGCAGCAAAACTAGATTTTCAATTTTGTGTTT 4200  
DB |||||  
QY 4201 TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAAACAGTTTCACTTGCAG 4260  
DB |||||  
QY 4201 TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAAACAGTTTCACTTGCAG 4260  
DB |||||

QY 4261 TGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTTCATATTTCCA 4320  
DB |||||  
QY 4261 TGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTTCATATTTCCA 4320  
DB |||||  
QY 4321 TATCCTGTCTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCCTCCGAAACTGGAAGTTATT 4380  
DB |||||  
QY 4321 TATCCTGTCTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCCTCCGAAACTGGAAGTTATT 4380  
DB |||||  
QY 4381 ATTTTAAATAAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440  
DB |||||  
QY 4381 ATTTTAAATAAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440  
DB |||||  
QY 4441 GATTCTTGTCTTATTACTGCTAAAAAATAAAAAA 4481  
DB |||||  
QY 4441 GATTCTTGTCTTATTACTGCTAAAAAATAAAAAA 4481  
DB |||||  
RESULT 6  
ADZ49164  
ID ADZ49164 standard; DNA; 4481 BP.  
XX  
AC ADZ49164;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Insulin signaling pathway related gene, SEQ ID 493.  
XX  
KW Antidiabetic; Gene Therapy; Non-insulin dependent diabetes;  
KW insulin resistance; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN US2005085436-A1.  
XX  
PD 21-APR-2005.  
XX  
PF 08-JUL-2004; 2004US-00887553.  
XX  
PR 08-JUL-2003; 2003US-0485883P.  
XX  
PA (LIHH/) LI H.  
PA (MAJU/) MA J.  
XX  
PI Li H, Ma J;  
XX  
WPI; 2005-305194/31.  
XX  
PT Treating, preventing or ameliorating pathological conditions associated  
PT with dysregulation of the insulin signaling pathway (ISP) comprises  
PT administering to a subject an amount of a modulator of any of the  
PT proteins regulated by ISP.  
XX  
PS Disclosure; SEQ ID NO 493; 70pp; English.  
XX  
CC The present invention relates to a method for treating, preventing or  
CC ameliorating pathological conditions associated with dysregulation of the  
CC insulin signaling pathway (ISP). The method comprises administering to a  
CC subject a modulator for ISP-regulated proteins or a pharmaceutical  
CC composition comprising the described modulator. The method is useful for  
CC treating, preventing or ameliorating pathological conditions associated  
CC with dysregulation of the ISP such as type II diabetes or type A syndrome  
CC of insulin resistance. The present sequence is a human homolog of a  
CC Drosophila gene regulated by ISP. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20050085436.  
XX  
SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

Query Match 100.0%; Score 4481; DB 14; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCCTCACTGTCTCCCTCCCGGCCCGGG 60  
DB |||||  
QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCCTCACTGTCTCCCTCCCGGCCCGGG 60  
DB |||||  
QY 61 GCCACCTCAGCTTCTGCTTCGCTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
DB |||||  
QY 61 GCCACCTCAGCTTCTGCTTCGCTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
DB |||||  
QY 121 CACCTCCGCTCCCAACCCGCGCTCCGCGCGCCCGCTCCGATCGCTCAAGCGCCGCA 180  
DB |||||  
QY 121 CACCTCCGCTCCCAACCCGCGCTCCGCGCGCCCGCTCCGATCGCTCAAGCGCCGCA 180  
DB |||||  
QY 181 GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCCAGTGGGCC 240  
DB |||||  
QY 181 GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCCAGTGGGCC 240  
DB |||||  
QY 241 GGGTGGCGCTCCAGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCG 300  
DB |||||  
QY 241 GGGTGGCGCTCCAGCGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCG 300  
DB |||||  
QY 301 CCGGACCGTATCCCTCCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
DB |||||  
QY 301 CCGGACCGTATCCCTCCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
DB |||||  
QY 361 CAGAGTCCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGGCAACCGCGCGCGCG 420  
DB |||||  
QY 361 CAGAGTCCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGGCAACCGCGCGCGCG 420  
DB |||||  
QY 421 TTTCCGCGCGCTCCCGCGCGCTTCTTGTCTCGCGCGCGCTCCCGCGCGCTTCTGTCGTC 480  
DB |||||  
QY 421 TTTCCGCGCGCTCCCGCGCGCTTCTTGTCTCGCGCGCGCTCCCGCGCGCTTCTGTCGTC 480  
DB |||||  
QY 481 GTCTTCTCCCTCCGCGCGCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB |||||  
QY 481 GTCTTCTCCCTCCGCGCGCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB |||||  
QY 541 CTCGCGCGCGCTGCTGCTCCCGCGCGCTTCCGCGCTTCTTGGCGCGCGCGCGCGCGCG 600  
DB |||||  
QY 541 CTCGCGCGCGCTGCTGCTCCCGCGCGCTTCCGCGCTTCTTGGCGCGCGCGCGCGCGCG 600  
DB |||||  
QY 601 TGTCTCCCGCGCGCTGCGAGCGGTGATGGCCCTCACCATGTCGCTGAAGCCCGAG 660  
DB |||||  
QY 601 TGTCTCCCGCGCGCTGCGAGCGGTGATGGCCCTCACCATGTCGCTGAAGCCCGAG 660  
DB |||||  
QY 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720  
DB |||||  
QY 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720  
DB |||||  
QY 721 CAGCCCGCGCGCGCTGCTCAATGTCGCAAGCCCGCGCGCAGCGCGCTTCTAGCGTCG 780  
DB |||||  
QY 721 CAGCCCGCGCGCGCTGCTCAATGTCGCAAGCCCGCGCGCAGCGCGCTTCTAGCGTCG 780  
DB |||||  
QY 781 CCGCGCGCGCGCTTCCGCGCTCCTCGTCTCGGTCTCTCTGCTCCGCGCAGCGCTCCC 840  
DB |||||  
QY 781 CCGCGCGCGCGCTTCCGCGCTCCTCGTCTCGGTCTCTCTGCTCCGCGCAGCGCTCCC 840  
DB |||||  
QY 841 TCCTCGGTGTCGCGCGACTCCGCGCGGAGCGCGCGCTCGGCGAGAGTCCGAAAC 900  
DB |||||  
QY 841 TCCTCGGTGTCGCGCGACTCCGCGCGGAGCGCGCGCTCGGCGAGAGTCCGAAAC 900  
DB |||||  
QY 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAATATGAG 960  
DB |||||  
QY 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAATATGAG 960  
DB |||||  
QY 961 ATGGTTTATATACATCAGTGTGGCTCCAAATGTGAAGTCAAGTGAAGTGA 1020  
DB |||||  
QY 961 ATGGTTTATATACATCAGTGTGGCTCCAAATGTGAAGTCAAGTGAAGTGA 1020  
DB |||||  
QY 1021 GGTATATACAGGAGTTTTTAAACTTACAGTCCGAGGTGATTTGGTACTTTGATGCC 1080  
DB |||||  
QY 1021 GGTATATACAGGAGTTTTTAAACTTACAGTCCGAGGTGATTTGGTACTTTGATGCC 1080  
DB |||||  
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCGGGCGCGAAACGCTGAAGAAATAATGGAGAGT 1140

DB |||||  
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCGGGCGCGAAACGCTGAAGAAATAATGGAGAGT 1140  
DB |||||  
QY 1141 ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATAGACTCCAGTTAT 1200  
DB |||||  
QY 1141 ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATAGACTCCAGTTAT 1200  
DB |||||  
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTCAATGGCGAACACAAA 1260  
DB |||||  
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTCAATGGCGAACACAAA 1260  
DB |||||  
QY 1261 GAGAAGGACCTGGAGCCCTGGGATGCGAGTGAATCAAGCAATAGAGAACTTGGAGCT 1320  
DB |||||  
QY 1261 GAGAAGGACCTGGAGCCCTGGGATGCGAGTGAATCAAGCAATAGAGAACTTGGAGCT 1320  
DB |||||  
QY 1321 TTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380  
DB |||||  
QY 1321 TTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380  
DB |||||  
QY 1381 AATTATGTTGATGCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440  
DB |||||  
QY 1381 AATTATGTTGATGCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440  
DB |||||  
QY 1441 AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGCAAAACAGTTAGCAAGAA 1500  
DB |||||  
QY 1441 AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGCAAAACAGTTAGCAAGAA 1500  
DB |||||  
QY 1501 ATTGAGTCAAGTGCCTCAGTCAAAAGCTCAGTGGCCCTGGAAAAATGATAGGAGTGA 1560  
DB |||||  
QY 1501 ATTGAGTCAAGTGCCTCAGTCAAAAGCTCAGTGGCCCTGGAAAAATGATAGGAGTGA 1560  
DB |||||  
QY 1561 GAAGAAAAATACAGAGCAGTTTCCAGTGAACGTGAGGGGCGACAGCATAAAC 1620  
DB |||||  
QY 1561 GAAGAAAAATACAGAGCAGTTTCCAGTGAACGTGAGGGGCGACAGCATAAAC 1620  
DB |||||  
QY 1621 ACTAGGAAAAATAATATATATCTCTCGACAAAGAAATAGAGAGTCTATATCTTGGGA 1680  
DB |||||  
QY 1621 ACTAGGAAAAATAATATATATCTCTCGGACAAAGAAATAGAGAGTCTATATCTTGGGA 1680  
DB |||||  
QY 1681 AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCCTGGATCGGGTCCATGCCATCAAGA 1740  
DB |||||  
QY 1681 AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCCTGGATCGGGTCCATGCCATCAAGA 1740  
DB |||||  
QY 1741 TCCACTTCTCAGCTTCAACCCGAAATTCCTGTTTCAAGCAAGAGAGTAGTTAAT 1800  
DB |||||  
QY 1741 TCCACTTCTCAGCTTCAACCCGAAATTCCTGTTTCAAGCAAGAGAGTAGTTAAT 1800  
DB |||||  
QY 1801 GGAGGTGTTCCCTGGCCATCGCTTCCCTCTCTCTCTCGCCACCTTCTCGCTAC 1860  
DB |||||  
QY 1801 GGAGGTGTTCCCTGGCCATCGCTTCCCTCTCTCTCTCGCCACCTTCTCGCTAC 1860  
DB |||||  
QY 1861 CAGTCAAGTCCCAACTCTCTTCCACTCGGGCAGCCACCCCTACAGCGCGCCCTCCAGG 1920  
DB |||||  
QY 1861 CAGTCAAGTCCCAACTCTCTTCCACTCGGGCAGCCACCCCTACAGCGCGCCCTCCAGG 1920  
DB |||||  
QY 1921 CCCCCCTCGCGCCATCCAGACCCCGCTCTACCCCTCTGCTCATGGTCTCCAGTCTCT 1980  
DB |||||  
QY 1921 CCCCCCTCGCGCCATCCAGACCCCGCTCTACCCCTCTGCTCATGGTCTCTCCAGTCTCT 1980  
DB |||||  
QY 1981 GTCTCTACTATGCTTAAACGCGATGTTTTCAGAGGGGCTCCAGAGGATGTCCTCCAAAGGCC 2040  
DB |||||  
QY 1981 GTCTCTACTATGCTTAAACGCGATGTTTTCAGAGGGGCTCCAGAGGATGTCCTCCAAAGGCC 2040  
DB |||||  
QY 2041 CAGCGACATCTTCGAAATCAGAGATTTCTGTGGGAGGGGTTTCCATATCCAGTGGCTA 2100  
DB |||||  
QY 2041 CAGCGACATCTTCGAAATCAGAGATTTCTGTGGGAGGGGTTTCCATATCCAGTGGCTA 2100  
DB |||||  
QY 2101 GAATTTGTATCCCAACCCAGCCAGTGAAGCAGTCTCTCTCCAGTAGCAAGACAGT 2160  
DB |||||  
QY 2101 GAATTTGTATCCCAACCCAGCCAGTGAAGCAGTCTCTCTCCAGTAGCAAGACAGT 2160  
DB |||||  
QY 2161 CCCTCGGGGGAGCGTGGTCTACAGTGGTTCAGTGGGTTCCAGAGTATTCCTCCCTAAACT 2220  
DB |||||

Db 2161 CCCTCGGGGGAACGTGTGTCATCAGTGTCACTGTGGGTTCCAAAGATTATCCCTTAAAACT 2220  
Qy 2221 CATAGACCCAGGTCTCCAGACAGAACAGTAGTTGGAATAACCCCAAGTGGGCGAGTTCTT 2280  
Db 2221 CATAGACCCAGGTCTCCAGACAGAACAGTAGTTGGAATAACCCCAAGTGGGCGAGTTCTT 2280  
Qy 2281 GCTTCTCCCAAGCTGGTATTATTCCAAACGTGAAGCTGTGGCATGCCATTCCAGCTGCA 2340  
Db 2281 GCTTCTCCCAAGCTGGTATTATTCCAAACGTGAAGCTGTGGCATGCCATTCCAGCTGCA 2340  
Qy 2341 TCTCCTAGCGCTGTAGTCCGTCATCGAACAGAGCTGTATCCCTTCTAGTGAGCTTAA 2400  
Db 2341 TCTCCTAGCGCTGTAGTCCGTCATCGAACAGAGCTGTATCCCTTCTAGTGAGCTTAA 2400  
Qy 2401 GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCCTCGAGGGAAATAAAGAAAATATTAAA 2460  
Db 2401 GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCCTCGAGGGAAATAAAGAAAATATTAAA 2460  
Qy 2461 CCCAATGAAACATCACCTAGCTTCTCAAAGCTGAAAAAACAAGGTATATCAACAGTTGTT 2520  
Db 2461 CCCAATGAAACATCACCTAGCTTCTCAAAGCTGAAAAAACAAGGTATATCAACAGTTGTT 2520  
Qy 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATTTAGGTTACAG 2580  
Db 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATTTAGGTTACAG 2580  
Qy 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAAACAAAAATAGAGAGGAGAAAA 2640  
Db 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAAACAAAAATAGAGAGGAGAAAA 2640  
Qy 2641 TCAAGAGATTGATCAAAAGACAAAATTGAACCAAGTGTCTAAGGATTTCTTCAATTGAAAT 2700  
Db 2641 TCAAGAGATTGATCAAAAGACAAAATTGAACCAAGTGTCTAAGGATTTCTTCAATTGAAAT 2700  
Qy 2701 AGCAGCAGCACTGTACCAAGTGGCAGCAGCAAGCCGAATAGCCCAAGATTTCCCTTCA 2760  
Db 2701 AGCAGCAGCACTGTACCAAGTGGCAGCAGCAAGCCGAATAGCCCAAGATTTCCCTTCA 2760  
Qy 2761 ATACTTAGTAAACACGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCCAGACT 2820  
Db 2761 ATACTTAGTAAACACGGAGCAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCCAGACT 2820  
Qy 2821 TCCAGCCCAGCATGTAAACAGAGAAAGACGATAGGAAAGAGAAAGACGACGTGAG 2880  
Db 2821 TCCAGCCCAGCATGTAAACAGAGAAAGACGATAGGAAAGAGAAAGACGACGTGAG 2880  
Qy 2881 CAAGTTAGGAATCAACATTGNAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCTCT 2940  
Db 2881 CAAGTTAGGAATCAACATTGNAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCTCT 2940  
Qy 2941 CAGCCAAAGCCCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCCATCT 3000  
Db 2941 CAGCCAAAGCCCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCCATCT 3000  
Qy 3001 ATGTGGGTGTCATCAACAGCCAACTCCAGTTTATCTCAGCCCTGTGTGTTTTGCACCAAT 3060  
Db 3001 ATGTGGGTGTCATCAACAGCCAACTCCAGTTTATCTCAGCCCTGTGTGTTTTGCACCAAT 3060  
Qy 3061 ATGATGTATCCAGTCCAGTGAGCCAGCGTGCAACCTTTATACCCAAATACCTATGACG 3120  
Db 3061 ATGATGTATCCAGTCCAGTGAGCCAGCGTGCAACCTTTATACCCAAATACCTATGACG 3120  
Qy 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAATATGCCCCACAGCGG 3180  
Db 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAATATGCCCCACAGCGG 3180  
Qy 3181 CAAGACCCAGCATCATCAGATGCCATGATGACCCAGCGGTGAGCAGCGGGCCACCGATT 3240  
Db 3181 CAAGACCCAGCATCATCAGATGCCATGATGACCCAGCGGTGAGCAGCGGGCCACCGATT 3240  
Qy 3241 GCAGCCACCCACAGCTTACTCAGCGCAATATGTTGCTTACAGTCCAGAGTTCCCA 3300  
Db 3241 GCAGCCACCCACAGCTTACTCAGCGCAATATGTTGCTTACAGTCCAGTCCAGAGTTCCCA 3300

Qy 3301 AATCAGGCCCTTGTTCAGCATGTGCCAATTTATCAGTCTCAGCATCTCTCATGTCTATAGT 3360  
Db 3301 AATCAGGCCCTTGTTCAGCATGTGCCAATTTATCAGTCTCAGCATCTCTCATGTCTATAGT 3360  
Qy 3361 CCTGTAATACAGGGTAATGCTTAGAATGATGGCAACACAAACACAGCCAGCCTGGTTTA 3420  
Db 3361 CCTGTAATACAGGGTAATGCTTAGAATGATGGCAACACAAACACAGCCAGCCTGGTTTA 3420  
Qy 3421 GTATCTTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCGCATGCGATGTATGCAAT 3480  
Db 3421 GTATCTTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCGCATGCGATGTATGCAAT 3480  
Qy 3481 CCCAAATTAACATACAAAGAGACAAAGCCCTTCTTTCTACTTTTGCCATTTCCACGGGC 3540  
Db 3481 CCCAAATTAACATACAAAGAGACAAAGCCCTTCTTTCTACTTTTGCCATTTCCACGGGC 3540  
Qy 3541 TCCCTTGTTCAGAGTAGTGGCAACCTTAAACGCTACCTTGCAACCCACATATCTCCACCCCT 3600  
Db 3541 TCCCTTGTTCAGAGTAGTGGCAACCTTAAACGCTACCTTGCAACCCACATATCTCCACCCCT 3600  
Qy 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCAAACATGTTGGAAGTCATCTTGCA 3660  
Db 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCAAACATGTTGGAAGTCATCTTGCA 3660  
Qy 3661 CCCAGTCTCTTCAGCAACCATCAGCACAGGCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720  
Db 3661 CCCAGTCTCTTCAGCAACCATCAGCACAGGCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720  
Qy 3721 CAGCAGCAGTCAGCCATTTTACCACGGGGGTTTCGGCAACTCCACCTTCATGACACCT 3780  
Db 3721 CAGCAGCAGTCAGCCATTTTACCACGGGGGTTTCGGCAACTCCACCTTCATGACACCT 3780  
Qy 3781 GCCTCCAAACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTACG 3840  
Db 3781 GCCTCCAAACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTACG 3840  
Qy 3841 ATCCATCTCTCTCAGCTTCAGCGCGGTATACCAACCCACCCACATGGCCACAGTACT 3900  
Db 3841 ATCCATCTCTCTCAGCTTCAGCGCGGTATACCAACCCACCCACATGGCCACAGTACT 3900  
Qy 3901 CAGCTCATGTACAGTCAGGAATGGTTCTTCTCATCAACTGCCCATGGCCCAATGATG 3960  
Db 3901 CAGCTCATGTACAGTCAGGAATGGTTCTTCTCATCAACTGCCCATGGCCCAATGATG 3960  
Qy 3961 CTAATGAGCAGACAGCCACCCGGCGGTCCCAGCGCCCTCGCTCAAAAGTGCATACAG 4020  
Db 3961 CTAATGAGCAGACAGCCACCCGGCGGTCCCAGCGCCCTCGCTCAAAAGTGCATACAG 4020  
Qy 4021 CCCATTCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGCAACCTTCAGTACAAGCC 4080  
Db 4021 CCCATTCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGCAACCTTCAGTACAAGCC 4080  
Qy 4081 CACCAACAAACAGCAGTTGTAAGGTGCGCTGGAGGAACCGAAAGCCAAATTCCTCTCTC 4140  
Db 4081 CACCAACAAACAGCAGTTGTAAGGTGCGCTGGAGGAACCGAAAGCCAAATTCCTCTCTC 4140  
Qy 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGAGAAAATAGNAATTTTCAATTTTGTGTTT 4200  
Db 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGAGAAAATAGNAATTTTCAATTTTGTGTTT 4200  
Qy 4201 TAAATATATATGTTGATTTTCTGTAAACATCAATAGGAATGCTAAACAGTTTCACTTGCG 4260  
Db 4201 TAAATATATATGTTGATTTTCTGTAAACATCAATAGGAATGCTAAACAGTTTCACTTGCG 4260  
Qy 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAATCTGGGGCTATTTCCATAATTTCCA 4320  
Db 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAATCTGGGGCTATTTCCATAATTTCCA 4320  
Qy 4321 TATGCTGTTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCCGAAACTGGAAGTTATTT 4380  
Db 4321 TATGCTGTTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCCGAAACTGGAAGTTATTT 4380



QY 1192 TCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGC 1251  
D5 1193 |||||  
D5 1081 TCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGC 1140  
QY 1252 GAAACAACAAAGAGAGAGACCTCGAGCCCTGGGATGAGGTGAACCTACAGACCAATGAGGAA 1311  
D5 1141 GAAACAACAAAGAGAGAGACCTCGAGCCCTGGGATGAGGTGAACCTACAGACCAATGAGGAA 1200  
QY 1312 CTTGAGGCTTTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATAT 1371  
D5 1201 CTTGAGGCTTTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATAT 1260  
QY 1372 AATGAGAAAATATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCTGTTATACAGTG 1431  
D5 1261 AATGAGAAAATATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCTGTTATACAGTG 1320  
QY 1432 CCCTTTAGAAAGAGATAACTCAGAGAAATTTTAAACCGGAAGCAAGGGCAAAACAGATT 1491  
D5 1321 CCCTTTAGAAAGAGATAACTCAGAGAAATTTTAAACCGGAAGCAAGGGCAAAACAGATT 1380  
QY 1492 GCAGAGAAAATTGAGTCAAGTGCCCAAGTACAAAGCTCGAGTGGCCCTGGAAAATGATGAT 1551  
D5 1381 GCAGAGAAAATTGAGTCAAGTGCCCAAGTACAAAGCTCGAGTGGCCCTGGAAAACGATGAT 1440  
QY 1552 AGGAGTAGGAGAAAATAACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCAC 1611  
D5 1441 AGGAGTAGGAGAAAATAACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCAC 1500  
QY 1612 AGCATAAACACTAGGAAAATAAATATATTTCTCTGGACAAAAGAAATAGAGAAAGTCA 1671  
D5 1501 AGCATAAACACTAGGAAAATAAATATATTTCTCTGGACAAAAGAAATAGAGAAAGTCA 1560  
QY 1672 TCCTGGGAAAGTGGGAGACAGAAATTCACCGGTATGGGCCAGCCCTGGATCGGGCTCCATG 1731  
D5 1561 TCCTGGGAAAGTGGGAGACAGAAATTCACCGGTATGGGCCAGCCCTGGATCGGGCTCCATG 1620  
QY 1732 CCATCAAGATCCATTCTCACACTTCAGATTTCAACCCGGAATTCGTGTTAGACCAAGA 1791  
D5 1621 CCATCAAGATCCATTCTCACACTTCAGATTTCAACCCGGAATTCGTGTTAGACCAAGA 1680  
QY 1792 GTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGGCCATCTCTCTCTCGCCCACT 1851  
D5 1681 GTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGGCCATCTCTCTCTCGCCCACT 1740  
QY 1852 TCTCGCTACAGTCAAGTCCCAACTCTCTTCCACTCGGGCAGCACCCCTACACGGCG 1911  
D5 1741 TCTCGCTACAGTCAAGTCCCAACTCTCTTCCACTCGGGCAGCACCCCTACACGGCG 1800  
QY 1912 CCCTCCAGGCCCTCGCGGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTGTTCT 1971  
D5 1801 CCCTCCAGGCCCTCGCGGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTGTTCT 1860  
QY 1972 CCAGTCTCTGCTCTACTATGCTTAAAGCGCATGCTTTCAGAAAGGCGCTTCAAGATGTC 2031  
D5 1861 CCAGTCTCTGCTCTACTATGCTTAAAGCGCATGCTTTCAGAAAGGCGCTTCAAGATGTC 1920  
QY 2032 CCAAGGCCCAAGCAGACATCTCGAAATCACAGATTTCTGCTGGGAGGGTTCATATCC 2091  
D5 1921 CCAAGGCCCAAGCAGACATCTCGAAATCACAGATTTCTGCTGGGAGGGTTCATATCC 1980  
QY 2092 AGTGGCTTAGAATTTGTATCCCAACACCCAGTGAAGCAGTATCTCTCCAGTAGCA 2151  
D5 1981 AGTGGCTTAGAATTTGTATCCCAACACCCAGTGAAGCAGTATCTCTCCAGTAGCA 2040  
QY 2152 AGGACCACTCTCGGGGGAAAGTGGTCAATCAGTGGTCAAGTGGTTCGAAGTATCC 2211  
D5 2041 AGGACCACTCTCGGGGGAAAGTGGTCAATCAGTGGTCAAGTGGTTCGAAGTATCC 2100  
QY 2212 CTTAAACCTATAGACCCAGGTCTCCACAGACAGAACTAGTATGGAAATPACCCCAAGTGG 2271  
D5 2101 CTTAAACCTATAGACCCAGGTCTCCACAGACAGAACTAGTATGGAAATPACCCCAAGTGG 2160  
QY 2272 CCAGTTCTTGTCTTCCCCAAAGCTGGTATATTTCAACTGAAGCTGTGGCCATGCTTAT 2331

D5 2161 CCAGTTCTTGTCTTCCCAAGCTGGTATATTTCAACTGAAGCTGTGGCCATGCTTAT 2220  
QY 2332 CCAGTGCATCTCTTACGCCCTGTAGTCTCGCATCGAACAGAGCTGTATACCCCTTCTAGT 2391  
D5 2221 CCAGTGCATCTCTTACGCCCTGTAGTCTCGCATCGAACAGAGCTGTATACCCCTTCTAGT 2280  
QY 2392 GAGCTAAAAGATTCACGGCTTCAAGATCAGAGGCAGAACTCTCTCGAGGGAAATAAGAA 2451  
D5 2281 GAGCTAAAAGATTCACGGCTTCAAGATCAGAGGCAGAACTCTCTCGAGGGAAATAAGAA 2340  
QY 2452 AATATTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAAGGTATATCA 2511  
D5 2341 AATATTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAAGGTATATCA 2400  
QY 2512 CCAGTTGTTTCTGAACTAGAAAACAGATTGATTTTAAAGAAATTTAAGAAATGATTTT 2571  
D5 2401 CCAGTTGTTTCTGAACTAGAAAACAGATTGATTTTAAAGAAATTTAAGAAATGATTTT 2460  
QY 2572 AGTTACAGCCAAAGTTCTTCTGAAATCTATGGATCAACTTACTATAACAAAATAGAGAG 2631  
D5 2461 AGTTACAGCCAAAGTTCTTCTGAAATCTATGGATCAACTTACTATAACAAAATAGAGAG 2520  
QY 2632 GGAGAAAATCAAGAGATTTGATCAAAAGCAAAATTCGAACTGCTAAGGATTTCTTC 2691  
D5 2521 GGAGAAAATCAAGAGATTTGATCAAAAGCAAAATTCGAACTGCTAAGGATTTCTTC 2580  
QY 2692 ATTCAAAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCAT 2751  
D5 2581 ATTCAAAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCAT 2640  
QY 2752 TCCCTTTCAATACTATTAGTAAACAGGACCAAGAGGGAACCTGAGGTCACTTCCCAAGG 2811  
D5 2641 TCCCTTTCAATACTATTAGTAAACAGGACCAAGAGGGAACCTGAGGTCACTTCCCAAGG 2700  
QY 2812 GTTCAGACTTCCAGCCAGCATGTAAACAAGAGAAAGCAGATAGGAAGAGAAAGAC 2871  
D5 2701 GTTCAGACTTCCAGCCAGCATGTAAACAAGAGAAAGCAGATAGGAAGAGAAAGAC 2760  
QY 2872 GCAGCTGAGCAAGTTAGGAAATCAACTTGAATCCCAATGCAAAAGGAGTTCAACCCAGCT 2931  
D5 2761 GCAGCTGAGCAAGTTAGGAAATCAACTTGAATCCCAATGCAAAAGGAGTTCAACCCAGCT 2820  
QY 2932 TCTTCTCTCAGCCAAAGCCCTTACTACCCCAACTTCACTCGGCCCTCAAGCAACCT 2991  
D5 2821 TCTTCTCTCAGCCAAAGCCCTTACTACCCCAACTTCACTCGGCCCTCAAGCAACCT 2880  
QY 2992 AGCCCATCTATGTTGGGTCTATCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTT 3051  
D5 2881 AGCCCATCTATGTTGGGTCTATCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTT 2940  
QY 3052 GCACCAATATGATGTTTCCAGTCCCAGTGAGCCAGCGTGCAACCTTTATACCCAATA 3111  
D5 2941 GCACCAATATGATGTTTCCAGTCCCAGTGAGCCAGCGTGCAACCTTTATACCCAATA 3000  
QY 3112 CCTATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCC 3171  
D5 3001 CCTATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCC 3060  
QY 3172 CAAAGCGGCAAGACAGCATCATCAGATGCCATGATGCACCAGCGTCAGACGCGGC 3231  
D5 3061 CAAAGCGGCAAGACAGCATCATCAGATGCCATGATGCACCAGCGTCAGACGCGGC 3120  
QY 3232 CCACCGATTCAGCCACCCACAGCTTACTTCCACGCAATATGTTGCTACAGTCCCTCAG 3291  
D5 3121 CCACCGATTCAGCCACCCACAGCTTACTTCCACGCAATATGTTGCTACAGTCCCTCAG 3180  
QY 3292 CAGTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCAATTATCATCAGTCTCAGCATCTCTCAT 3351  
D5 3181 CAGTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCAATTATCATCAGTCTCAGCATCTCTCAT 3240  
QY 3352 GTCTATGTTCTGTAAATACAGGGTAAATGCTAGATGATGGCACCAACCAACACGCGCAG 3411



DR WPI; 1997-281034/25.  
DR P-PSDB; AAW24800, AAW24801.  
XX  
PT Antibody 1C2 used for treating or preventing neuro-degenerative diseases  
PT - associated with proteins containing long poly:glutamine repeats, e.g.  
PT Huntington's disease.  
XX  
PS Claim 21; Page 45-47; 69pp; French.  
XX  
CC The invention relates to a monoclonal antibody (Mab) 1C2 for the  
CC treatment of neurodegenerative diseases associated with the presence of  
CC polyglutamine repeat regions. This Mab is already known for its affinity  
CC to the TAR4 binding protein (TBP) transcription initiation factor,  
CC especially at the amino acid sequence LEEQRQOQQQ found at the N-  
CC terminus of TBP. Mab 1C2 has been shown to have a high affinity for  
CC polyglutamine repeats with a proportional affinity to the number of  
CC glutamine repeats. This affinity has been used to identify genes encoding  
CC proteins containing long polyglutamine repeats which are implicated in  
CC neurodegenerative diseases. A screen of an expression library, generated  
CC from a lymphoblastic cell line from a patient suffering from  
CC spinocerebellar ataxia (SCA), with Mab 1C2 isolated 6 new sequences  
CC (AAAT78906-T78911) encoding polyglutamine repeats. Mab 1C2 also isolated  
CC the complete SCA2 gene in clone DAN1 (sequence presented here). The  
CC sequence appears to contain 2 open reading frames (ORF) the second of  
CC which may be generated by an frameshift slippage or by an alternative  
CC splicing event. The first ORF also encodes a 22 amino acid polyglutamine  
CC repeat region near the N-terminus of the protein. Normal SCA2 alleles  
CC contain 17-29 CAG triplet repeats with 1-3 CAA repeats interspersed  
CC whereas the mutant sequence from patients with SCA contains at least 30,  
CC preferably 37-50 CAG repeats. Mab 1C2, active fragment of it or nucleic  
CC acids encoding it are specifically used to treat Huntington's disease,  
CC SCA types 1-5 or 7, X-linked spino-bulbar muscular atrophy (Kennedy  
CC disease), dentarorubral-pallidolusial atrophy, dominant autosomal  
CC spinocerebellar ataxia, familial spastic paraplegia, bipolar affective  
CC disorder, manic depressive psychoses and schizophrenia  
XX  
SQ Sequence 4200 BP; 1152 A; 1200 C; 913 G; 935 T; 0 U; 0 Other;  
Query Match 88.1%; Score 3949; DB 2; Length 4200;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 4020; Conservative 0; Mismatches 0; Indels 61; Gaps 1;  
QY 401 ACGGCAACGGCGGCGCGGTTTCGGCCCGGCTCCGGCGGCTCTCGGTCGCGGCGG 460  
DB 1 ACGGCAACGGCGGCGCGGTTTCGGCCCGGCTCCGGCGGCTCTCGGTCGCGGCGG 60  
QY 461 CTCGCCGCGCCCTTCGTGCTGCTTCCTCCCTCGCCAGCCCGGCGCCCTCGGCGG 520  
DB 61 CTCGCCGCGCCCTTCGTGCTGCTTCCTCCCTCGCCAGCCCGGCGCCCTCGGCGG 120  
QY 521 CGCCAAACCGCGCCTCCCGGCTCGGCGCCGCTGCGTCCCGCGGTTCCGCGGCTCTCCT 580  
DB 121 CGCCAAACCGCGCCTCCCGGCTCGGCGCCGCTGCGTCCCGCGGTTCCGCGGCTCTCCT 180  
QY 581 TGGCGCGCGCGCTCCCGGCTGTCGCCCGCGCGCTGCGAGCGGTGATGGGCCCTCA 640  
DB 181 TGGCGCGCGCGCTCCCGGCTGTCGCCCGCGCGCTGCGAGCGGTGATGGGCCCTCA 240  
QY 641 CCATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 700  
DB 241 CCATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 300  
QY 701 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 760  
DB 301 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 360  
QY 761 GCAGCGGCTTCTAGCTGTCGCCCGCGCGGCTTCGCCGCTGCTGCTGCTGCTGCT 820  
DB 361 GCAGCGGCTTCTAGCTGTCGCCCGCGCGGCTTCGCCGCTGCTGCTGCTGCTGCT 420  
QY 821 GGTCTCGGCGCAGGCTCCCTCTCGGTGCTGCGGCGGCTTCGCCGCGGCGGCGGCGG 880  
DB 421 GGTCTCGGCGCAGGCTCCCTCTCGGTGCTGCGGCGGCTTCGCCGCGGCGGCGGCGG 480

QY 881 GCCTGGGCGAGGTCGAAACAGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGATG 940  
DB 481 GCCTGGGCGAGGTCGAAACAGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGATG 540  
QY 941 GAATCTATGCAAAATATGAGGATGGTTTCATATACCTTACATCAGTTGTTGGCTCCAAATGTG 1000  
DB 541 GAATCTATGCAAAATATGAGGATGGTTTCATATACCTTACATCAGTTGTTGGCTCCAAATGTG 600  
QY 1001 AAGTACAAAGTCAAAATATGAGGATGGTTTCATATGAGGATGGTTTCATATGAGGATGGTT 1060  
DB 601 AAGTACAAAGTCAAAATATGAGGATGGTTTCATATGAGGATGGTTTCATATGAGGATGGTT 660  
QY 1061 GTGATTTGGTACTCTTGTATGCGCCACATCAGAAAAGTACAGAAATCCAGTTTCGGGGCCGAAAC 1120  
DB 661 GTGATTTGGTACTCTTGTATGCGCCACATCAGAAAAGTACAGAAATCCAGTTTCGGGGCCGAAAC 720  
QY 1121 GTGAAGAAATATGAGGATGGTTTCATATGAGGATGGTTTCATATGAGGATGGTTTCATATGAGGATGGTT 1180  
DB 721 GTGAAGAAATATGAGGATGGTTTCATATGAGGATGGTTTCATATGAGGATGGTTTCATATGAGGATGGTT 780  
QY 1181 AAGATATGAGTCCAGTTTCATATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTA 1240  
DB 781 AAGATATGAGTCCAGTTTCATATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTA 840  
QY 1241 AAGTGAATGGCGAACACAAAGAGAGAACCTGGAGCCCTGGGATGCGAGGTGAATCAGTACAG 1300  
DB 841 AAGTGAATGGCGAACACAAAGAGAGAACCTGGAGCCCTGGGATGCGAGGTGAATCAGTACAG 900  
QY 1301 CCAATGAGGAATCTGAGGCTTTGGAAAATGACGTTCTAATGGATGGGATCCCAATGATA 1360  
DB 901 CCAATGAGGAATCTGAGGCTTTGGAAAATGACGTTCTAATGGATGGGATCCCAATGATA 960  
QY 1361 TGTTTCGATATAATGAAGAAAATTTATGCTGAGTGTCTACGTTATGATAGAGATTTATCTT 1420  
DB 961 TGTTTCGATATAATGAAGAAAATTTATGCTGAGTGTCTACGTTATGATAGCAGTTTATCTT 1020  
QY 1421 CGTATACAGTGCCTTTAGAAAAGAGATAACTCAGAAAGAAATTTTTAAACGGGAAGCAAGG 1480  
DB 1021 CGTATACAGTGCCTTTAGAAAAGAGATAACTCAGAAAGAAATTTTTAAACGGGAAGCAAGG 1080  
QY 1481 CAAACCAAGTTAGCAGAGAAATTTGAGTCAAGTGCAGTGCAGTCAAAAGCTCAGTGGCCCTGG 1540  
DB 1081 CAAACCAAGTTAGCAGAGAAATTTGAGTCAAGTGCAGTGCAGTCAAAAGCTCAGTGGCCCTGG 1140  
QY 1541 AAAATGATGATAGGAGTGAGGAAGAAAATACACAGCAGTTTCAGAGAAATTTCCAGTGAAC 1600  
DB 1141 AAAATGATGATAGGAGTGAGGAAGAAAATACACAGCAGTTTCAGAGAAATTTCCAGTGAAC 1200  
QY 1601 GTGAGGGGACAGCAGATAAACAATAAGGAAAATAAATATATTTCTCTGGACAAAGAAAATA 1660  
DB 1201 GTGAGGGGACAGCAGATAAACAATAAGGAAAATAAATATATTTCTCTGGACAAAGAAAATA 1260  
QY 1661 GAGAAGTCATATCTTGGGAAAGTGGGAGACAGAAATTCACCGCGTATGGGCCAGCGCTGGAT 1720  
DB 1261 GAGAAGTCATATCTTGGGAAAGTGGGAGACAGAAATTCACCGCGTATGGGCCAGCGCTGGAT 1320  
QY 1721 CGGCTCCATGCCATCAAGATCCACTTCTCACCTTCAGATTTTCAACCCGAAATTTCTGCTT 1780  
DB 1321 CGGCTCCATGCCATCAAGATCCACTTCTCACCTTCAGATTTTCAACCCGAAATTTCTGCTT 1380  
QY 1781 CAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCCATCTCTCTTCT 1840  
DB 1381 CAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCCATCTCTCTTCTCT 1440  
QY 1841 CTCGCCACCTTCTCGCTACCGTCCAGTCCCAACTCTCTTCCAGCTCGGGCAGGCCACC 1900  
DB 1441 CTCGCCACCTTCTCGCTACCGTCCAGTCCCAACTCTCTTCCAGCTCGGGCAGGCCACC 1500  
QY 1901 CTACACGGCGCCCTCCAGGCGCCCTCGGGCCCATCAGACCCCGCTCTCACCCCTCTG 1960  
DB 1501 CTACACGGCGCCCTCCAGGCGCCCTCGGGCCCATCAGACCCCGCTCTCACCCCTCTG 1560







Db 3660 AAAGCCAAATCCCTCCTCCTCCTACTGCTTCTACCAACTGGAAGCACAGAAACTAG 3719  
Qy 4181 AATTTCATTATTTTGGTTTTTAAATAATATATGTTGATTTCCTGTAACATCAATAGCAA 4240  
Db 3720 AATTTCATTATTTTGGTTTTTAAATAATATATGTTGATTTCCTGTAACATCAATAGCAA 3779  
Qy 4241 TGCTAACAGTTCACCTGAGTGGGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACCTG 4300  
Db 3780 TGCTAACAGTTCACCTGAGTGGGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACCTG 3839  
Qy 4301 GGGGTATTCCATATTCATATGCTGTTTCAGAGTCCCGAGGTPACCCAGCTCTGCTT 4360  
Db 3840 GGGGTATTCCATATTCATATGCTGTTTCAGAGTCCCGAGGTPACCCAGCTCTGCTT 3899  
Qy 4361 GCCGAACTGGAAGTATTTATTTTAAATACCTTGAAAGTTCATGAACACATCAGCTA 4420  
Db 3900 GCCGAACTGGAAGTATTTATTTTAAATACCTTGAAAGTTCATGAACACATCAGCTA 3959  
Qy 4421 GCAAAAGAGTAACAAAGAGTGATTCCTGCTGCTATTACTGCTAATAAAAAA 4480  
Db 3960 GCAAAAGAGTAACAAAGAGTGATTCCTGCTGCTATTACTGCTAATAAAAAA 4019  
Qy 4481 A 4481  
Db 4020 A 4020

RESULT 9  
ID ABV75290 standard; DNA; 4225 BP.  
XX AC ABV75290;  
XX DT 07-MAR-2003 (first entry)  
XX DE Mouse SCA2 polypeptide encoding DNA.  
XX KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;  
KW memory impairment; mutant; mouse; gene; ds.  
XX OS Mus musculus.  
FH Key Location/Qualifiers  
FT CDS 27..3884  
FT /\*tag= a  
FT /transl\_except= (pos: 222..224, aa: Xaa)  
FT /note= "Xaa is Pro"  
FT /product= "SCA2"  
FN WO200289567-A2.  
XX PD 14-NOV-2002.  
XX PF 07-MAY-2002; 2002WO-US014698.  
XX PR 07-MAY-2001; 2001US-0289231P.  
XX PA (CEDA-) CEDARS SINAI MEDICAL CENT.  
XX PI Pulst SM;  
XX DR WPI; 2003-103467/09.  
XX DR P-PSDB; ABB82697.  
XX PT New mutant mouse, useful as a disease model for identifying agents for  
PT treating obesity or memory impairment, comprises a disrupted  
PT spinocerebellar ataxia 2 gene.  
XX FS Disclosure; Page 56-62; 76pp; English.  
XX CC The invention relates to a mutant non-human mammal having a disrupted  
CC spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a  
CC disease model for identifying agents or compounds that may treat obesity

CC or memory impairment. The stem cell is used for generating the mutant  
CC animal. The present sequence represents the nucleotide sequence (GenBank  
XX Accession No. AF041472) encoding the mouse SCA2 polypeptide  
SQ Sequence 4225 BP; 1007 A; 1324 C; 1042 G; 851 T; 0 U; 1 Other;  
Query Match 70.0%; Score 3138.2; DB 10; Length 4225;  
Best Local Similarity 86.2%; Pred. No. 0;  
Matches 3662; Conservative 1; Mismatches 473; Indels 112; Gaps 14;  
Qy 222 CGCGCAGCAGGTGGCCCGGGTGGCGCTCGCTCAGGGCGGGCGGGAGCGGGCG 281  
Db 71 CCCGAGCGCGCGCGCGGGCTGGCGCGCGGGCTCGCTCGCGAGCGGGCGCG 130  
Qy 282 GGGCGGGGTGGCGCGCGCGCGCGACCGTATCCCTCCGCCGCCCTCCCGCGCGGGCC 341  
Db 131 GCGCGGGGGGGGGCGGGCTGGGTATCCCTCCGCCGCCCTCCCGCGGGCGCC 190  
Qy 342 CGGCCCCCTCCCTCCCGCAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGTAGCAA 401  
Db 191 CGGCGCCCTCCCGCGGGCGCGCTCGCCACCSTGGCCCTCAGACTGTTTGTAGCAA 250  
Qy 402 CGGCAACGGCGGGCGGGGTTCGGCGCGGCTCCGGCGGGCTCCTTGGTCTCGGGGGCC 461  
Db 251 CGGCCA-----CGGCGGCTCCCGCGCGGGCTCCCGCGGGCTGCTCGGGGCC 304  
Qy 462 TCCCGCGCCCTTCGTGCTGCTTCTCCCTCCCGCAGCGCGGGCGCGCTCCGGCGCG 521  
Db 305 TCCCGCGCCCTTCGTGCTGCTTCTGCTGCTCTG-----GCCCGCGGGCGCAC 352  
Qy 522 GCCAACCGCGCCCTCCCGCTCGCGCGCGGTGCGTCCCGCGCGGTTCGGCGTCTCCTT 581  
Db 353 GCGCGCGCGCGCTGCGCGCGCGGTCCGGCGGTCCCGCGCGCTCCCGCGTCTCCTC 412  
Qy 582 ---GGGGCGCGCGCTCCCGCTGCTCCCGCGCGGTGCGTCCCGCGCGGTTCGGCGTCT 638  
Db 413 CTCGGCGCGCGCGCACCGGCTGCTCCCGCGCGCGGTGCGAGCGGGTATGGGCGCT 472  
Qy 639 CACCATGTGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCA 698  
Db 473 CACCATGTGCTGAAGC-----  
Qy 699 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 758  
Db 491 -----GCAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 532  
Qy 759 CGGCAGCGGCTTCTAGGTCGCGCGCGCGCGCGCTTCGGCGTCTCGTCTCGGTCTC 818  
Db 533 CGGCGGGCTGCTCTGTCGTCGCGCGCGCGCGCGCGCTTCGGCGGTGACCTCGGGCTC 592  
Qy 819 CTCGTCCTCGGCCACGGCTCCTCCTCGGTG-----GTGCGCGCGCACCTCCGGCGGGAG 875  
Db 593 CGTGGTGGCGCGCGCGCGCGCGGTGGCGTCTTCCTCGCGCGCGCGGGCGGGCG 652  
Qy 876 GCCCGGCTGGGCAGAGGTGAAACAGTAACAAAGGACTGCTCAGTCTACGATTTCTTT 935  
Db 653 TCCCGGCTGGGCAGAGGTGGAACAGTAGCAAGGACTGCTCAGCTCAGATTTCTTT 712  
Qy 936 TGATGGAATCTATGCAAAATATGAGATGGTTTCATATACATCAGTTGTTGGTCCAA 995  
Db 713 TGATGGAATCTATGCAAAACGTGAGGATGGTTTCATATACATCAGTTGTTGGATCGAA 772  
Qy 996 ATGTGAAGTACAAAGTGAATAATGAGGTATATATGAAGGAGTTTTTAAACCTTACAGTCC 1055  
Db 773 ATGTGAAGTACAAAGTGAATAACGAGGCATATATGAAGGAGTTTTTAAACATACAGTCC 832  
Qy 1056 GAAGTGTGATTTGGTACTTGTATGCCGACATGAGAAAAGTACAGAATCCAGTTCGGGGCC 1115  
Db 833 TAAGTGTGACTTGGTACTTGTATGTCGACATGAGAAAAGTACAGAATCCAGTTCGGGGCC 892  
Qy 1116 GAAACGTGAAGAAATATGAGAGTATTTGTTTCAAAATGTTACAGACTTTGTTGGTACA 1175  
Db 893 AAAACGTGAAGAAATATGAGAGTATTTGTTTCAAAATGCTCAGACTTCGTTGTTGGTACA 952





QY 885 GGGCAGAGGTCAAACAGTAAACAAAGGAGCTGCTCAGTCTACGATTTCTTTTGATGGAAT 944  
DB |||||  
QY 235 GGGCAGAGTCCGAAACAGTAGCAAGGAGCTGCTCAGCTACGATTTCTTTTGATGGAAT 294  
DB |||||  
QY 945 CTATCAAAATAGAGGATGTTTCATATACATCAGTTGTGGCTCCAAATGGAAGT 1004  
DB |||||  
QY 295 CTATGCAACGTGAGGATGTTTCATATACATCAGTTGTGGATCGAAATGGAAGT 354  
DB |||||  
QY 1005 ACAAGTGAATAATGGAGGTATATGAAGGAGTTTTTAAATCTACAGTCCGAAAGTGA 1064  
DB |||||  
QY 355 ACAAGTGAATAACCGAGGCATATGAAGGAGTTTTTAAACATACAGTCTCAAGTGA 414  
DB |||||  
QY 1065 TTTGGTACTTGTATGCGGCACATGAGAAAGTACAGATCCAGTTGCGGCGCCGAAACGTGA 1124  
DB |||||  
QY 415 CTTGGTACTTGTATGTCACATGAGAAAGTACAGATCCAGTTGCGGCGCCGAAACGTGA 474  
DB |||||  
QY 1125 AGAAATATGGAGAGTATTTTCTCAATGTTTCAGACTTTGTGGTACAGTTTAAAGA 1184  
DB |||||  
QY 475 AGAAATATGGAGAGTATTTTCTCAATGCTCAGACTTCGTTGGTACAGTTTAAAGA 534  
DB |||||  
QY 1185 TATGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGT 1244  
DB |||||  
QY 535 TACAGACTCCAGTTATGACGAGGAGATGCTTTTACTGACTCTGCTCTCAGCGCAAGT 594  
DB |||||  
QY 1245 GAATGCGCAACAAGAGAGGAGCTGGAGCCCTGGGATGAGTGAACTCACAGCCAA 1304  
DB |||||  
QY 595 GAATGCTGAGCAACAGGAGAGGAGCTGGAGCCCTGGGATGAGTGAGGAGCTCACGCGCAG 654  
DB |||||  
QY 1305 TGAGGAACTTGAGGCTTTTGGAAATGACGTATCTAATGATGGATCCCAATGATATGTT 1364  
DB |||||  
QY 655 CGAGGAGCTGGAG---CTGGAGAAATGATGTCTAATGGATGGAGCCCAATGACATGTT 711  
DB |||||  
QY 1365 TCGATATAATGAAGAAATATGTTGCTAGTGTCTAGTATGATAGCAGTTTATCTTCGTA 1424  
DB |||||  
QY 712 TCGATATAATGAAGAAATATGTTGCTAGTGTCTAGTATGATAGCAGTTTATCTTCATA 771  
DB |||||  
QY 1425 TACAGTGCCTTTAGAAAGAGATAAATCAGAGAAATTTTTTAAACGGGAGAGCAAGGCAAA 1484  
DB |||||  
QY 772 TACGGTTCTTTAGAAAGGAGCAACTCAGAGAAATTTCTTAAACGGGAGGCAAGGCAAA 831  
DB |||||  
QY 1485 CCAGTTAGCAGAGAAATTTGAGTCAAGTCCAGTACAAAGCTCCAGTGGCCCTCGAATA 1544  
DB |||||  
QY 832 CCAGTTAGCAGAGAAATTTGAAATCCAGTGTCTCAGTACAAAGCTCGTGTGCGCCCTTGAGAA 891  
DB |||||  
QY 1545 TGAATGATAGAGTGAGGAGAAATAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGA 1604  
DB |||||  
QY 892 TGAATGACCGAGTGAGGAGAAATAATACACAGCAGTCCAGAGAACTGAGTGACCGGGA 951  
DB |||||  
QY 1605 GGGGCA CAGCATAAACACTAGGGAATAATATATTTCTCTCGGCAAAAGAAATAGAGA 1664  
DB |||||  
QY 952 GGGGCATGCCCCCAACACTAGGAGCAATAAATATATTTCTCTCGGCAAAAGAAACAGAGA 1011  
DB |||||  
QY 1665 AGTCATATCTGGGGAAGTGAGGAGACAGAAATTCACCGGTATGGGCGAGCCTGGATCGGG 1724  
DB |||||  
QY 1012 AGTCCTATCTCTGGGGAAGTGAGGAGACAGAGCTCACACGGATGGGCGCAGCTGGGCGCAG 1071  
DB |||||  
QY 1725 CTTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTTCAACCCGAAATTCGTTTCAGA 1784  
DB |||||  
QY 1072 CTCATGCGCTCAAGAGTGTCTCTCACACTTCAGATTTTCAACCCGAAACGCTGGCTCAGA 1131  
DB |||||  
QY 1785 CCAAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCCTGGCCATCTCTCTCTCTCG 1844  
DB |||||  
QY 1132 CCAAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCCTGGCCATCTCTCTCTCTCG 1191  
DB |||||  
QY 1845 CCACACTTCTCGTACAGTTCAGTTCGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1904  
DB |||||  
QY 1192 CCACACTTCT 1251  
DB |||||  
QY 1905 ACGGCGCCCTCCAGGCGCCCTCTCGGCGCCATCCAGACCCCGCTCTCACCCCTCTCTCA 1964  
DB |||||  
QY 1252 ACGGCGCCCTCCAGGCGCCCTCTCGAGGCCATCCAGACCCCGCTCTCACCCCTCTCTCA 1311  
DB |||||

QY 1965 TGGTTCTCCAGCTCTGTCTCTTACTATGCTTAAACGCAATGCTTTCAGAGGCGCTCCAAAG 2024  
DB |||||  
QY 1312 TGGTTCTCCAGCTCTGTCTCTTACTATGCTTAAACGCAATGCTTTCAGAGGAGCCCAAG 1371  
DB |||||  
QY 2025 GATGTCCCAAAAGGCCAGCGACATCTCTGAAATCACAGAGTTTCTGTGGGAGGGTTC 2084  
DB |||||  
QY 1372 GATGTCTCAAAAGGCACAGCGCCACCTCTCGAAATCACAGAGTCTCTGTGGGAGGGTTC 1431  
DB |||||  
QY 2085 CATATCCAGTGGCTAGAAATTTGTATCCCAAAACCCAGTGAAGCAGTCTCTCTCC 2144  
DB |||||  
QY 1432 CATGTCTAGTGGCTAGAAATTTGTATCCCAAAATCCCAAGTGAAGCAGTCTCTCTCC 1491  
DB |||||  
QY 2145 AGTAGCAGGACCAAGTCCCTCGGGGGAAACGTGTCTATCAGTGTCTAGTGGGTTCCAAAG 2204  
DB |||||  
QY 1492 AGTGGCAAGGACCAAGTCTCTCAGGGGGAAACGTGTCTCAGTGTCTAGTGGGTTCCAAAG 1551  
DB |||||  
QY 2205 ATTATCCCTCTAAACCTCATAGACCCAGGTCTCCACAGACAGAAACAGTATTTGAAATACCC 2264  
DB |||||  
QY 1552 GTTATCTCCAAACTCACAGACCCAGGTCTCCAGGCAGAGCAGCATTTGAAACTCTCC 1611  
DB |||||  
QY 2265 CAGTGGCCAGTCTTGTCTCTCCCAAGCTGGTATTTATTTCCAACTGAAGCTGTGGCCAT 2324  
DB |||||  
QY 1612 CAGCGGCTGTGTCTTGTCTCTCCCAAGCTGGCATCATCTCCAGAGAGCGTCTTCCAT 1671  
DB |||||  
QY 2325 GCCTATTTCCAGTGTCTCTTACGCTGTAGTCTCTGCATCGAACAGAGTGTTCACCC 2384  
DB |||||  
QY 1672 GCCTGTCTCCCGCGCATCTCCGACTCTCTGCCAGCCCTGCATCTCAACAGAGCACTGACCCC 1731  
DB |||||  
QY 2385 TTCTAGTGGGCTAAAGATTTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCTGCAAGGAA 2444  
DB |||||  
QY 1732 ATCTATTTGAGGCAAAAGATTTCCAGGCTTCAAGATCAGAGGCAAGTCTCTCTGCAAGGAG 1791  
DB |||||  
QY 2445 TAAAGAAATATTAACCCCAATCACTAGCTTCTCAAAAGCTGAAACAAAGG 2504  
DB |||||  
QY 1792 TAAAGAAATGTTAAAGCAAGTGAACATCACCTAGCTTTTCAAAAGCTGCAACAAAGG 1851  
DB |||||  
QY 2505 TATATCACAGTGTCTTCTGAACATAGAAACAGATTTGATGATTTTAAAGAAATTTAAGAA 2564  
DB |||||  
QY 1852 TATGTCCACAGTGTCTGTAACACAGAAACAGATTTGATGACTTAAAGAAATTTAAGAA 1911  
DB |||||  
QY 2565 TGATTTTAGTTTACAGCCAAAGTCTTACTTCTCAATCTATGGATCAACTACTAAACAAAA 2624  
DB |||||  
QY 1912 TGATTTTAGTTTACAGCCAAAGTCTTACATCTGAATCTATGGATCAACTACTAAGCAAAA 1971  
DB |||||  
QY 2625 TAGAGAGGAGAAATCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTGTAAAGA 2684  
DB |||||  
QY 1972 TAGAGAGGAGAAAGTCAAGAGATTTGATTTAAGATAAAACGGAAGCAAGTGTAAAGA 2031  
DB |||||  
QY 2685 TTCCTTTCAITGA-----AAATAGCAGCAGCAAACTGTACAGTGGCAGCAGCAAGCCGAA 2738  
DB |||||  
QY 2032 TAGTTTTCATGACAGCAGCAGCAGCAAGTGTACAGTGGCAGCAGCAAGCCAA 2091  
DB |||||  
QY 2739 TAGCCCCAGCATTTCCCTTCAATCTTAGTAAACGGAGCACAGAGGGGACCTCAGGT 2798  
DB |||||  
QY 2092 CAGCCCTAGACTCTCCCTTCCATGCTTAGTAAATGACAGAGCAAGAGGGGCTCAGGT 2151  
DB |||||  
QY 2799 CACTTTCCCAAGGGGTTTCAGACTTCCAGCCCAAGTGTAAACAAAGAGAAACAGATAAAGA 2858  
DB |||||  
QY 2152 CACATCCCAAGGGTGCAGACTTCCAGCCCAAGCTGCAAAACAAAGAGAGATGACAGAGA 2211  
DB |||||  
QY 2859 AGAGAAAGAACGCGCTGAGCAAGTTAGGAAATCAACATTTGATCCCAATGCAAGGA 2918  
DB |||||  
QY 2212 AGAGAAAGAACACACAGAGCAGGTTAGGAAATCGACATTTGAAATCCCAATGCAAGGA 2271  
DB |||||  
QY 2919 GTTCAACCCAGCTCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTTCACCTCGGCC 2978  
DB |||||  
QY 2272 GTTCAACCCCTCGTCTCTCTCAGCCAAAGCTTCTACTACCCCAAGCTCAGCTCGGCC 2331  
DB |||||  
QY 2979 TCAAGCACAACCTTAGCCCATCTATGTTGGGTTCATCAACAGCCCAACTTCCAGTTTATCTCA 3038  
DB |||||  
QY 2332 TCAAGCACAACCCAGCCCATCTATGTTGGGTTCATCAGCAGCCAGCTCCAGTGTACACTCA 2391  
DB |||||  
QY 3039 GCCTGTTGTTTTCACCAAAATATGATGTATCCAGTCCCAGTGAGCCCGGCTGCAACC 3098  
DB |||||

	Db	2392	 GCCTGTGTTGGCACCCAAATAGTATCGCGTCCTCAGTGAGCGCGGCGTAGCAACC	2451
	Qy	3099	T T T A T A C C C A A T A C T T A G A C G C C A T G C C A G T G A A T C A A G C C A A G A C A T A T A G A G C - - -	3155
	Db	2452	T T T A T A C C C A A T A C T T A G A C G C C A T G C C T G T G N A C C A A G A C C A G A C A T A T A G A G C A G G	2511
	Qy	3156	- - - A G T A C C A A A T A T G C C C C A A C A G C G G C A A G A C C A G C A T C A T C A G A G T G C C A T G A T G C A	3212
	Db	2512	T A A A G T A C C A A A T A T G C C C C A A C A G C A A G A C C A A C A T C A T C A A A G A C C A C A T G A T G C A	2571
	Qy	3213	C C C A G C G T C A G C A G C G G G C C C A C G A T T G C A G C C A C C C A C C A G C T T T A C T C A C G C A A T A	3272
	Db	2572	C C C A G C C T C C G C G G C A G G G C C A C C A T G T A G C C A C C C G C C G C T T A C T C C A C T C A G T A	2631
	Qy	3273	T G T T G C C C T A C A G T C C T C A G C A G T T C C C A A A T C A G C C C C T T G T T C A G C A T G T G C C A C A T T A	3332
	Db	2632	C G T T G C C T A C A G C C C T C A G C A G T T C C A A T C A G C C T T T G T C A G A T G T G C G C A T T A	2691
	Qy	3333	T C A G T C T C A G A C T C C T C A T G T C T A T A G T C C T G T A A T A C A G G T A A T G C T A G A A T G A T G G C	3392
	Db	2692	T C A G T C T C A G A C C T C A T G T A C A G C T C T G T C A T A C A A G S T A A T G C C A G G A T G A T G G C	2751
	Qy	3393	A C C A C A A C A C A C G C C C A G C C T G T T T A G T A T C T T C T T C A G A A C T C A G T A C C G G G C T C A	3452
	Db	2752	A C C A C A G C A C A T G C T C A G C C T G S T T F A G T G T C T T C T T C A G C T G C T C A G T T C G S G G C T C A	2811
	Qy	3453	T G A C A C A C A T G C G A T G T A T G A T C T G C T C C A A A T T A C A T A C A A A A G A G A C A A G C C C	3512
	Db	2812	C G A C A G A C G A C G C C A T G A T G C A A U T C C A A T T A C C A T A C A A A G A G A C A A G C C C	2871
	Qy	3513	T T C T T T C T A C T T T G C C A T T T C A C G G G C T C C C T T G C T C A G C A G T A T G C G A C C C T A A C G C	3572
	Db	2872	T T C T T T C T A C T T T G C A T T T C C A C G G C T C C C T C G C T C A G C A G T A T G C A C A T C C T A A T G C	2931
	Qy	3573	T A C C T T G C A C C C A T A C T C C A C A C C C T A G C C T T A G T T A C C C C A C T C G A C A G A C A G C A	3632
	Db	2932	C G C C C T G C A T C C A C A T A C T C C C A T C C T C A G C C T T C G S C A C T C C C A C C G G A C A G A C A	2991
	Qy	3633	A A G C C A C A T G G T G G A A G T C A T C T G C A C C A C C A G T C H G T T C A G C A C C A T C A G C A C C A G C	3692
	Db	2992	A A G C C A C A T G G T G G A A G T C A C C C T G C A C C A G T C T G T T C A G C A C C A T C A G C A C C A G C	3051
	Qy	3693	C G C C A G G C T C T C C A T C T G G C C A G T C C A C A G C A G C A G T C A G C C A T T T A C C A C G G G G C T	3752
	Db	3052	T G C C A G G C T T C A T C T G G C C A G T C C A C A G C A G C A G T C G G C C A T T T A T C A T G G G G G C T	3111
	Qy	3753	T G G C C A A C T C C A C C C T C C A T A C A C T G C C T C C A A C C G A G T C G C C A C A G A A T A G T T T	3812
	Db	3112	G G C A C C A A C C A C C T T C C A T G A C A C C T G C C T C T A A T A C A C A G T C T C C A C A G A G C A G T T	3171
	Qy	3813	C C C A G A G C A C A C A G A C T G C T T T A G A T C C A T C T T C T C A G T T C A G C G C G G G G T A T A C	3872
	Db	3172	C C C A G A G C A A C A G A C A G T C T T C C A C A T C C A C C C T T C T A T G T T C A G C G C G C A T A C A C	3231
	Qy	3873	C A A C C C A C C C A C A T G C C C A C G T A C C T C A G G C T C A T G T A C A G T C A G A A T G T T C C T T C	3932
	Db	3232	C A C C C A C C C A C A T G C C C A C A G T A C C T C A G B C T C A T G T A C A G T C A G A A T G T T C C T T C	3291
	Qy	3933	T C A T C C A A C T G C C C A T G C G C C A A T G A T G C T A A T A G A C A C A G C C A C C C G G C G G T C C C C A	3992
	Db	3292	T C A T C C A A C T G C C C A T G C G C C A A T A T G C T A A T A G A C A C A G C A C C - - - C G S T C C C A A	3348
	Qy	3993	G G C G C C C T C G C T C A A A G T G C A T A C A G C C A T T C A G T C A G A C A C A G C C A C C C G G C G G T C C C C A	4052
	Db	3349	G G C G C C C T C G C T C A A A G T G C A C T A C A G C C C A T T C C A G C C A T T C C A G T T C G A C A C A C G C A T T T C C C	3408
	Qy	4053	C T A T A T A C A C C C C T T C A G T A C A A G C C C A C A C C A C A C A G C A G T T G T A A G S C T G C C C T G G	4112
	Db	3409	T T A T A T A G A C G A C C C T T C A G T A C A G C C C A C C A C A C A G A G T T G T A A G S C T G C C T T G G	3468
	Qy	4113	A G A A C C G A A G G C C A A A T T C C C T C C T C C C T T A C T G C T T C T A C C A A C T G G A A G C A C A G	4172

## RESULT 11

REGD L  
AAF21628

AAZ21628  
ID AAF21628 standard: DNA: 2006 BP.

XX  
DT  
07012 JHY

AC AAF21628;

XX  
CONTENTS

DT 27-MAR-2001 (first entry)

[illegible]

DE Human breast and ovarian c

XX

KW Human; breast cancer; ovar

KW nootropic; neuroprotective

KW antidiabetic; antiinflamm

KW antibacterial; antifungal;

KW Addison's disease; allergy

KW autoimmune thyroiditis; di

KW multiple sclerosis; rheuma

KW cardiovascular disorder; wv

XX  
50  
Homo sapiens

US  
yy  
Homo sapiens.

AA  
PN  
W0200055173-A1

FN W0200053173-A1.  
XX

21-SEP-2000.

XX  
07-DEC-2000

08-MAR-2000: 2000WO-US00058

[illegible]

PR 12-MAR-1999; 99US-012427

XX  
XX  
-----

PA (HUMA-) HUMAN GENOME SCI I

XX

PI Rosen CA, Ruben SM;

XX

DR WPI; 2000-611515/58.

DR P-PSDB; AAB58725.

XX

PT New human breast and ovar







KW osteoporosis; neurological disease; Alzheimer's disease;  
 KW Parkinson's disease; dementia; short memory; cancer;  
 KW sense or motor function; emotional reaction; fear response; panic;  
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
 KW tranquiliser.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1447413-A2.  
 XX  
 XX 18-AUG-2004.  
 XX  
 PF 12-FEB-2004; 2004EP-00003145.  
 XX  
 XX 14-FEB-2003; 2003JP-00102207.  
 PR 09-MAY-2003; 2003JP-00131452.  
 XX  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
 XX  
 DR WPI: 2004-583265/57.  
 DR P-PSDB; ADRI0292.  
 XX  
 PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
 XX  
 PS Claim 1; SEQ ID NO 1842; 2686pp; English.  
 XX  
 CC This invention relates to novel, isolated full length human cDNA  
 CC molecules and the encoded proteins thereof. Specifically, it refers to  
 CC cDNA clones obtained by an oligo-capping method, where none of these  
 CC clones are identical to any known human mRNAs. The present invention  
 CC describes an immunoassay to identify agonists and antagonists, as well as  
 CC antibodies, antisense molecules and siRNAs that can all be used to bind  
 CC to and modulate expression of the cDNA molecules. As such, these  
 CC molecules are useful for diagnostic markers or therapeutic targets for  
 CC the various diseases or morbid states. In particular, they are useful in  
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
 CC disease, Parkinson's disease, dementia, short memory and various cancers,  
 CC as well as for maintaining equilibrium of sense or motor function, and  
 CC for treating emotional reaction, fear response and panic. Accordingly,  
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
 CC cytostatic and tranquiliser activities. This polynucleotide is a full  
 CC length human cDNA sequence of the invention. NOTE: This sequence is not  
 CC given in the sequence listing of the specification but can be obtained on  
 CC CD-ROM from the European Patent Office, Vienna Sub-office.  
 XX  
 SQ Sequence 4369 BP; 1047 A; 1122 C; 1051 G; 1149 T; 0 U; 0 Other;  
 XX  
 Query Match 14.1%; Score 633; DB 13; Length 4369;  
 Best Local Similarity 92.8%; Pred. No. 2.7e-112;  
 Matches 697; Conservative 0; Mismatches 0; Indels 54; Gaps 1;  
 QY 3154 GCAGTACCAATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGCATGATGCAC 3213  
 DB 96 GCAGTACCAATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGCATGATGCAC 155  
 QY 3214 CCAGCGTCAGAGGGGGCCCGCCGATTCAGCCACCCAGCTTACTCCAGCAATAT 3273  
 DB 156 CCAGCGTCAGAGGGGGCCCGCCGATTCAGCCACCCAGCTTACTCCAGCAATAT 215  
 QY 3274 GTTGCTTACAGTCTCTCAGCAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCACATTAT 3333  
 DB 216 GTTGCTTACAGTCTCTCAGCAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCACATTAT 275  
 QY 3334 CAGTCTCAGCATCTCTCATGTCTATAGTCCTTAATACAGGGTAATGCTAGAATGATGGCA 3393  
 DB 276 CAGTCTCAGCATCTCTCATGTCTATAGTCCTTAATACAGGGTAATGCTAGAATGATGGCA 335  
 QY 3394 CCACCAACACAGCGCGGCTGGTTAGTATCTTTCAGCAACTCAGTACGGGGCTCAT 3453  
 XX  
 DB 336 CCACCAACACAGCGCGGCTGGTTAGTATCTTTCAGCAACTCAGTACGGGGCTCAT 395  
 QY 3454 GAGCAGAGCGATGCGATGTATGTCATGTCCCAAAATTACATACAAAGAGACAAGCCCT 3513  
 DB 396 GAGCAGAGCGATGCGATGTATG----- 417  
 QY 3514 TCTTTTCTACTTTGGCCATTTTCAGAGGGCTCCCTTGTGCTCAGCAGTATGCGCACCTTAACGCT 3573  
 DB 418 -----TTTCCACAGGGCTCCCTTGTGCTCAGCAGTATGCGCACCTTAACGCT 461  
 QY 3574 ACCCTGACCCACATATCTCCACACCCCTCAGCCTTTCAGCTTACCCCTTGGACAGAGCAA 3633  
 DB 462 ACCCTGACCCACATATCTCCACACCCCTCAGCCTTTCAGCTTACCCCTTGGACAGAGCAA 521  
 QY 3634 AGCCAAACATGTGGAAAGTCATCTCCACCCAGTCTCTGTTTTCAGCAGCATCAGCACAGGCC 3693  
 DB 522 AGCCAAACATGTGGAAAGTCATCTCTCCACCCAGTCTCTGTTTTCAGCAGCATCAGCACAGGCC 581  
 QY 3694 GCCCAGGCTCTCCATCTTCGSCAGTCCACAGCAGCAGTCCAGCCATTTACCACCGGGGCTT 3753  
 DB 582 GCCCAGGCTCTCCATCTTCGSCAGTCCACAGCAGCAGTCCAGCCATTTACCACCGGGGCTT 641  
 QY 3754 GCGCCAACTCCACCCCTCCATGACACCTGCTCCAAACAGCAGTCCGACAGCAATAGTTTC 3813  
 DB 642 GCGCCAACTCCACCCCTCCATGACACCTGCTCCAAACAGCAGTCCGACAGCAATAGTTTC 701  
 QY 3814 CCAGCAGCAACAGCAGTCTTTTACGATCCATCTTTCAGCTTTCAGCGCGGCTATACC 3873  
 DB 702 CCAGCAGCAACAGCAGTCTTTTACGATCCATCTTTCAGCTTTCAGCGCGGCTATACC 761  
 QY 3874 AACCCACCCACATGCGCCAGCTTACCTCAGG 3904  
 DB 762 AACCCACCCACATGCGCCAGCTTACCTCAGG 792  
 RESULT 14  
 AAV17229  
 ID AAV17229 standard; DNA; 623 BP.  
 AC AAV17229;  
 DT 29-JUN-1998 (first entry)  
 XX SCA2 gene fragment.  
 DE SCA2 gene; spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 CDS 341..583  
 FT /\*tag= a  
 FT /note= "SCA2 protein fragment, no stop codon given"  
 XX  
 PN WO9803679-A1.  
 XX  
 XX 29-JAN-1998.  
 PD  
 XX  
 PF 18-JUL-1996; 96WO-JP001999.  
 XX  
 XX 18-JUL-1996; 96WO-JP001999.  
 PR  
 XX  
 PA (SRLS-) SRL INC.  
 XX  
 XX Teuji S, Sanpei K;  
 PI  
 XX WPI; 1998-120796/11.  
 DR P-PSDB; AAW41372.  
 XX  
 XX Diagnosing spinocerebellar ataxis type II - by PCR and determining number  
 PT of CAG repeat units.  
 XX  
 XX Example 1; Page 11-12; 23pp; Japanese.  
 PS



[illegible]

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 17:46:01 ; Search time 15479 Seconds  
(without alignments)  
13544.345 Million cell updates/sec

Title: US-10-802-228-1  
Perfect score: 4481  
Sequence: 1 acccccggaggaagcaaccca.....taaaaaaaaaaaaaaaaaaaaa 4481

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1327	29.6	1528	4	CR619107 full-leng
2	903	20.2	912	5	BQ883680 AGENCOURT
3	899.8	20.1	1000	5	EX402110 EX402110
4	899	20.1	1060	1	AL554896 AL554896
5	897	20.0	938	5	EX375946 EX375946
6	850.6	19.0	977	7	CN646951 ILLUMIGEN
7	842	18.8	1100	3	BM455214 AGENCOURT
8	822.8	18.4	859	5	BQ716397 AGENCOURT
9	819.2	18.3	889	6	CA489164 AGENCOURT
10	797.6	17.8	846	1	AU124606 AU124606
11	785.2	17.5	1066	1	AL554863 AL554863
12	773.6	17.3	863	6	CD616049 CD616049
13	767	17.1	830	6	CD616048 CD616048
14	761.2	17.0	877	7	CO886647 BovGen.14
15	753.2	16.8	784	6	CA775370 i888f10.x
16	747	16.7	750	1	AU142284 AU142284
17	744.6	16.6	940	5	BU159977 AGENCOURT
18	739.6	16.5	877	5	BQ428363 AGENCOURT
19	736.6	16.4	856	7	CJ039644 CJ039644
20	736	16.4	878	6	CD106563 AGENCOURT
21	732.6	16.3	779	1	AU121620 AU121620
22	729.8	16.3	804	2	BG201697 RST21039

ALIGNMENTS

RESULT 1  
CR619107  
LOCUS full-length cDNA clone CS0D1087Y117 of Placenta Cot 25-normalized  
DEFINITION of Homo sapiens (human)  
ACCESSION CR619107.1 GI:50499914  
VERSION CR619107.1  
KEYWORDS HTC; CNSLT cDNA.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1528)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
Genoscope.  
2 (bases 1 to 1528)  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
FEATURES Location/Qualifiers  
source 1..1528  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1087Y117"  
/tissue\_type="Placenta Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

ORIGIN

Query Match 29.6%; Score 1327; DB 4; Length 1528;  
Best Local Similarity 89.7%; Pred. No. 3.5e-292;  
Matches 1522; Conservative 0; Mismatches 0; Indels 175; Gaps 2;

QY	2743	CCAGCATTTTCCCTTCAATACTTAGTAAACGAGACACAGAGGGACCTTAGGTCACCT	2803
DB	1	CCAGCATTTTCCCTTCAATACTTAGTAAACGAGACACAGAGGGACCTTAGGTCACCT	60
QY	2803	TCCCAGGGGTTTCAGACTTCAGGCCAGCATGTAAACAAGAGAAGACGATNAGGAAG	2862
DB	61	TCCCAGGGGTTTCAGACTTCAGGCCAGCATGTAAACAAGAGAAGACGATNAGGAAG	120
QY	2863	AAGAAAGACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAGGAGTTC	2922
DB	121	AAGAAAGACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAGGAGTTC	180
QY	2923	AACCCAGCTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTTCGSCCTCAA	2982
DB	181	AACCCAGCTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTTCGSCCTCAA	240
QY	2983	GCACAACCTAGCCCATCTATGGTGGGTCAACAACGCCAATCTCAGTTTATATCTCAGCCT	3042
DB	241	GCACAACCTAGCCCATCTATGGTGGGTCAACAACGCCAATCTCAGTTTATATCTCAGCCT	300
QY	3043	GTTTGTTGTTGCAACCAATATGATGTATCCAGTCCCAAGTGCCCAAGCGCTGCAACCTTTA	3102
DB	301	GTTTGTTGTTGCAACCAATATGATGTATCCAGTCCCAAGTGCCCAAGCGCTGCAACCTTTA	360
QY	3103	TACCCAATACCTPATGACGCCCATGCCAGTGAATCAAGCCAAGACATATAGAGC-----A	3156
DB	361	TACCCAATACCTPATGACGCCCATGCCAGTGAATCAAGCCAAGACATATAGAGCAGGTAAA	420
QY	3157	GTACCAATATGCCCAACAGCGGCAAGACCAAGCATCATCAGATGCCATATGACCCCA	3216
DB	421	GTACCAATATATGCCCAACAGCGGCAAGACCAAGCATCATCAGATGCCATATGACCCCA	480
QY	3217	GGGTCAAGCAGCGGGCCCAACCGATTTCAGCGCAACCCACAGCTTACTCCAGCAATATGTT	3276
DB	481	GGGTCAAGCAGCGGGCCCAACCGATTTCAGCGCAACCCACAGCTTACTCCAGCAATATGTT	540
QY	3277	GCCTACAGTCTCAGCAGTTCCCAATCAGGCCCTTGTTTACGATGTGCCACATTATCAG	3336
DB	541	GCCTACAGTCTCAGCAGTTCCCAATCAGGCCCTTGTTTACGATGTGCCACATTATCAG	600
QY	3337	TCTCAGCATCTCATGTCTATAGTCTGTATATACAGGGTAATGCTAGATATGGCAACCA	3396
DB	601	TCTCAGCATCTCATGTCTATAGTCTGTATATACAGGGTAATGCTAGATATGGCAACCA	660
QY	3397	CCAAACACACGCCCGCTGGTTTACTATCTTCTTTCAGCAACTCAGTACGGGGCTCATGAG	3456
DB	661	CCAAACACACGCCCGCTGGTTTACTATCTTCTTTCAGCAACTCAGTACGGGGCTCATGAG	720
QY	3457	CAGACGATGCGATGTATGCAATGCCAAATTTACATAACAAGGAGACAAGCCCTTCT	3516
DB	721	CAGACGATGCGATGTATGCAATGCCAAATTTACATAACAAGGAGACAAGCCCTTCT	780
QY	3517	TTCTACTTTGCCATTTTCCAGGGGCTCCTTGCTCAGCAAGTATGGCAACCTTAAGCTTACC	3576
DB	781	TTCTACTTTGCCATTTTCCAGGGGCTCCTTGCTCAGCAAGTATGGCAACCTTAAGCTTACC	840
QY	3577	CTGCAACCCATATCTCACACCTCAGCGCTTTCAGTACCCCACTGCAACGACGACCAAGC	3636
DB	841	CTGCAACCCATATCTCACACCTCAGCGCTTTCAGTACCCCACTGCAACGACGACCAAGC	900
QY	3637	CAACATGGTGGAGTCATCTTGCAACCAAGTCTGTTTACGACCACTCAGCACCGCCGCC	3696
DB	901	CAACATGGTGGAGTCATCTTGCAACCAAGTCTGTTTACGACCACTCAGCACCGCCGCC	960
QY	3697	CAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCATTTTACACGCGGGCTTGGC	3756
DB	961	CAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCATTTTACACGCGGGCTTGGC	1020
QY	3757	CCAACTCCACCTCCATGCACTTCCCTTCCAAACGAGTCCGACCAAGATAGTTTCCCA	3816
DB	1021	CCAACTCCACCTCCATGCACTTCCCTTCCAAACGAGTCCGACCAAGATAGTTTCCCA	1080
QY	3817	GCAGCACACAGACTGTCTTTACGATCCATCTTCTCAGCTTACGCGGGGTATACCAAC	3876

[illegible]

```
source
1. .912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6179729"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupeki dorsal root ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCACGCGTCCG-3' and
5'-GACTAGTTCTTAGTCGCGAGCGCCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupeki, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies.."
```

ORIGIN

Query Match 20.2%; Score 903; DB 5; Length 912;  
Best Local Similarity 99.3%; Pred. No. 3.4e-195;  
Matches 906; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2281 GCTTCTCCCAAGCTGGTATTATTCACACTGAAGCTGTGGCCATGCTATTCAGCTGCA 2340  
DB 1 GCTTCTCCCAAGCTGGTATTATTCACACTGAAGCTGTGGCCATGCTATTCAGCTGCA 60  
QY 2341 TCTCTACGGCTCTAGTCTGATCGACAGAGCTGTACCCCTCTAGTGAGGTAA 2400  
DB 61 TCTCTACGGCTCTAGTCTGATCGACAGAGCTGTACCCCTCTAGTGAGGTAA 120  
QY 2401 GATTTCAGGCTTCAAGATCAGAGCGAGACTCTCTCGAGGGAATAAGAAATATTTAA 2460  
DB 121 GATTTCAGGCTTCAAGATCAGAGCGAGACTCTCTCGAGGGAATAAGAAATATTTAA 180  
QY 2461 CCAATGAACATCAGCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAAGTTGT 2520  
DB 181 CCAATGAACATCAGCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAAGTTGT 240  
QY 2521 TCTGAACATGAACACAGATTGATGATTTAAGAAATTTAAGATGATTTAGTTACAG 2580  
DB 241 TCTGAACATGAACACAGATTGATGATTTAAGAAATTTAAGATGATTTAGTTACAG 300  
QY 2581 CCAAGTTCTACTTCTGATCTATGGATCAACTACTAAACAAATAAGAGGAGAGAA 2640  
DB 301 CCAAGTTCTACTTCTGATCTATGGATCAACTACTAAACAAATAAGAGGAGAGAA 360  
QY 2641 TCAAGAGATTGTATCAAGACAAATTTGAACCAAGTCTAAGGATTTCTTTCATTGAAAT 2700  
DB 361 TCAAGAGATTGTATCAAGACAAATTTGAACCAAGTCTAAGGATTTCTTTCATTGAAAT 420  
QY 2701 AGCAGCGCACTGTACCACTGGCAGCAGCAGCAGCGCAATAGCCCGCATTTCCCTTCA 2760  
DB 421 AGCAGCGCACTGTACCACTGGCAGCAGCAGCGCAATAGCCCGCATTTCCCTTCA 480  
QY 2761 ATACTTAGTAACGCGAGCAGAGAGGAGCTGAGTCACTTCCCAAGGGTTTCAGACT 2820  
DB 481 ATACTTAGTAACGCGAGCAGAGAGGAGCTGAGTCACTTCCCAAGGGTTTCAGACT 540  
QY 2821 TCCAGCCCGCAGCATGTAACAGAGAAAGACGATGAAGGAAGAGAAAGACGCGCTGAG 2880  
DB 541 TCCAGCCCGCAGCATGTAACAGAGAAAGACGATGAAGGAAGAGAAAGACGCGCTGAG 600  
QY 2881 CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGATTCAACCCAGCTTCTTCTCT 2940  
DB 601 CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGATTCAACCCAGCTTCTTCTCT 660  
QY 2941 CAGCCAAAGCTTCTACTACCCCACTTCACTCGGCTCAAGCAGACACCTAGCCCATCT 3000  
DB 661 CAGCCAAAGCTTCTACTACCCCACTTCACTCGGCTCAAGCAGACACCTAGCCCATCT 720

QY 3001 ATGGTGGGTGCATCAACAGCCAACTCCAGTTTATCTAGCTGCTTTGTTTTCACCAAT 3060  
DB 721 ATGGTGGGTGCATCAACAGCCAACTCCAGTTTATCTAGCTGCTTTGTTTTCACCAAT 780  
QY 3061 ATGATGTATCCAGTCCCGAGCCAGCGGTGCAACCTTTTATACCAATACCTATGACG 3120  
DB 781 ATGATGTATCCAGTCCCGAGCCAGCGGTGCAACCTTTTATACCAATACCTATGACG 840  
QY 3121 CCATGCCAGTGAATCAAGCCAAAGACATATAGAGAGATACCAATATGCCCAACAGCGG 3180  
DB 841 CCATGCCAGTGAATCCAGCCAAAGACATATAGAGAGATACCAATATGCCCAACAGCGG 900  
QY 3181 CAAGACACGAT 3192  
DB 901 CAAGACACGAT 912

RESULT 3  
BX402110 1000 bp mRNA linear EST 01-MAY-2004  
LOCUS BX402110 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
DEFINITION CDNA clone CSODC015B14 5-PRIME, mRNA sequence.  
ACCESSION BX402110  
VERSION BX402110.2 GI:46922451  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1000)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 13, 2003 this sequence version replaced gi:30632041.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
6451.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS1AC004ZF08QPlk=6451.r.

FEATURES  
source  
1. .1000  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODC015B14"  
/issue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 20.1%; Score 899.8; DB 5; Length 1000;  
Best Local Similarity 96.7%; Pred. No. 1.9e-194;  
Matches 927; Conservative 10; Mismatches 18; Indels 4; Gaps 2;

QY 327 TCCCCCGCGCGCCCGCCCTCCCTCCCGCAGAGCTCCCTCCCTCCCGCTCAGA 386  
DB 1 TCCCCGGATCGCGCCCGCCCTCCCTCCCGCAGAGCTCCCTCCCTCCCGCTCAGA 60  
QY 387 CTGTTTGTAGTGAACAGGCAACCGCGCGCGGTTCGGCCCGGCTCCCGGGGCTCT 446  
DB 61 CTGTTTGTAGTGAACAGGCAACCGCGCGGTTCGGCCCGGCTCCCGGGGCTCT 120



```
QY 447 TGGTCTCGGCGGCTCCCGCCCTTCTGTCGTCTCTTCTCCCTCCCTCGCCGCGG 506
Db 121 TGGTCTCGGCGGCTDCCGCGCTTCTGTCGTCTCTTCTCCCTCCCTCGCCGCGG 180
QY 507 CGCCCTCGGCGGCGCAACCGGCGCTCCCGCTCGGCGCGCTGGTTCGCCGCGG 566
Db 181 CGCCCTCGGCGGCGCAACCGGCGCTCCCGCTCGGCGCGCTGGTTCGCCGCGG 240
QY 567 TTCGCGGCTCTCTTGGGCGGCGCGGCTCCCGCTGTCGCCGCGGCGTGGAGCCGCT 626
Db 241 NTCCGCGGCTCTCTTGGGCGGCGCGGCTCCCGCTGTCGCCGCGGCGTGGAGCCGCT 300
QY 627 GTATGGGCGCTCAACATGTCTGCTGAAGCCCGAGCAGCAGCAGCAGCAACA 686
Db 301 GTATGGGCGCTCAACATGTCTGCTGAAGCCCGAGCAGCAGCAGCAGCAACA 360
QY 687 GCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 743
Db 361 GCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 420
QY 744 TGTCCGCAAGCCCGCGGCGAGCGG-CCTTCTAGCGTCCCGCGCGCGCTTCGCCGT 802
Db 421 TGTCCGCAAGCCCGCGGCGAGCGGCTTCTAGCGTCCCGCGCGCGCTTCGCCGT 480
QY 803 CCTCGTCTCGGTCTCTCGCTCGGCGAGGCTCCCTCTCGGTGTCGCGGCGACCT 862
Db 481 CCTCGTCTCGGTCTCTCGCTCGGCGAGGCTCCCTCTCGGTGTCGCGGCGACCT 540
QY 863 CGGCGGCGGAGCGCGCGCTGGCAGAGGTGCGAAACAGTAACAAAGACTGCTCAGT 922
Db 541 CGGCGGCGGAGCGCGCGCTGGCAGAGGTGCGAAACAGTAACAAAGACTGCTCAGT 600
QY 923 CTACGATTTCTTTTGTGGAATCTATGCAATATGAGATGTTTCTATATCTTACATCAG 982
Db 601 CTACGATTTCTTTTGTGGAATCTATGCAATATGAGATGTTTCTATATCTTACATCAG 660
QY 983 TTGTTGGCTCCAAATGTGAAGTACAAAGTCAAAATGGAGGTATATATGAAGAGTTT 1042
Db 661 TTGTTGGCTCCAAATGTGAAGTACAAAGTCAAAATGGAGGTATATATGAAGAGTTT 720
QY 1043 AAACCTACAGTCCGAAGTGTGATTTGGTACTTGATGCGGCACATGAGAAAGTACAGAA 1102
Db 721 AAACCTACAGTCCGAAGTGTGATTTGGTACTTGATGCGGCACATGAGAAAGTACAGAA 780
QY 1103 CCAGTTCGGGCGGCAACGCTGAAGAAATAATGGAGAGTATTTGTTCAAAATGTTCCAGACT 1162
Db 781 CCAGTTCGGGCGGCAACGCTGAAGAAATAATGGAGAGTATTTGTTCAAAATGTTCCAGACT 840
QY 1163 TTGTTGGTACAGTTTAAAGATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTG 1222
Db 841 TTGTTGGTACAGTTTAAAGATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTG 900
QY 1223 ACTCTGCTACAGTCTAAAGTGAATGCGCAACAAAGAGAGAGACCTGGAGCCCTGG 1281
Db 901 ACTCTGCTACAGTCTAAAGTGAATGCGCAACAAAGAGAGAGACCTGGAGCCCTGGG 959

RESULT 4
AL554896 1060 bp mRNA linear EST 30-MAR-2004
LOCUS AL554896 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI087Y117 5-PRIME, mRNA sequence.
ACCESSION AL554896
VERSION AL554896.3 GI:45859645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1060)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
```

```
TITLE
JOURNAL
COMMENT
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31276706.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6451.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI087AE09QPI&c=6451.r.
FEATURES
source
1..1060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI087Y117"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized."
ORIGIN
Query Match 20.1%; Score 899; DB 1; Length 1060;
Best Local Similarity 97.3%; Pred. No. 2.9e-194;
Matches 941; Conservative 6; Mismatches 12; Indels 8; Gaps 3;
QY 2743 CCCAGCATTTCCCTTCAATACTTAGTAACAGCGAGCACAAGAGGGACCTGAGGTCACT 2802
Db 1 CCCAGCATTTCCCTTCAATACTTAGTAACAGCGAGCACAAGAGGGACCTGAGGTCACT 60
QY 2803 TCCCAAGGGGTTTCCAGACTTCCAGCGCCAGCATGTAAAACAAGAGAAGACGATAAAGAAAG 2862
Db 61 TCCCAAGGGGTTTCCAGACTTCCAGCGCCAGCATGTAAAACAAGAGAAGACGATAAAGAAAG 120
QY 2863 AAGAAAGACGAGCTGAGCAAGTTAGGAAATCAACATTGAAATCCCAATGCAAAAGAGTTC 2922
Db 121 AAGAAAGACGAGCTGAGCAAGTTAGGAAATCAACATTGAAATCCCAATGCAAAAGAGTTC 180
QY 2923 AACCCAGTTCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTCCACTCGGCTCAA 2982
Db 181 AACCCAGTTCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTCCACTCGGCTCAA 240
QY 2983 GCACAACTAGCCCATCTATGGTGGGTCTATCAACAGCCCAACTCCAGATTTTACTCAGCCT 3042
Db 241 GCACAACTAGCCCATCTATGGTGGGTCTATCAACAGCCCAACTCCAGATTTTACTCAGCCT 300
QY 3043 GTTTGTTTGGACCAAAATATGATGTATCCAGTCCAGTGAGCCGAGCGGTGCAACCTTTTA 3102
Db 301 GTTTGTTTGGACCAAAATATGATGTATCCAGTCCAGTGAGCCGAGCGGTGCAACCTTTTA 360
QY 3103 TACCAATACCTATGAGCCCATCGCAGTGAATCAAGCAAGACATATAGAGC-----A 3156
Db 361 TACCAATACCTATGAGCCCATCGCAGTGAATCAAGCAAGACATATAGAGAGGTTAAA 420
QY 3157 GTACCAAAATATGCCCAACAGCGCAAGACCCAGCATCATCAGATGGCATATGATGACCCCA 3216
Db 421 GTACCAAAATATGCCCAACAGCGCAAGACCCAGCATCATCAGATGGCATATGATGACCCCA 480
QY 3217 GCGTCAGAGCGGCGCCACCGATTGTCAGCCACCCACAGCTTACTCCACCAATATGTT 3276
Db 481 GCGTCA-CAGCGGCGCCACCGATTGTCAGCCACCCACAGCTTACTCCACCAATATGTT 539
QY 3277 GCCTTACAGTCTCTCAGCAGTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAG 3336
Db 540 GCCTTACAGTCTCTCAGCAGTACCCAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAG 599
```



[illegible]

Db	841	TTTTACTGACTCTGCTATCAGTGGCTTAAGTGAATGCGGACCCCAAGAGAGAGGACTGG	900
QY	1274	A--GCCCTGGTGATGAGGTGAA---CTCACAGCCAAATGAGGAACCTTTGAGCGTTTGAAAA	1328
Db	901	NAGCCCTGGGGATGCGAGGGGAACCTCCAGCCCAATAGGGAACCTTTGAAGCTTTTGGNAA	960
RESULT 8			
BQ716397			
LOCUS			
DEFINITION	BQ716397	859 bp mRNA linear EST 16-JUL-2002	
ACCESSION	IMAGE:6196657	5', mRNA sequence.	
VERSION	BQ716397		
KEYWORDS	BQ716397.1	GI:21855294	
SOURCE			
ORGANISM		Homo sapiens (human)	
		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
		Homnidae; Homo.	
REFERENCE		1 (bases 1 to 859)	
AUTHORS		NIH-MGC <a href="http://imgc.nci.nih.gov/">http://imgc.nci.nih.gov/</a> .	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs@mail.nih.gov">cgapbs@mail.nih.gov</a> CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM13605 row: h column: 02 High quality sequence stop: 659. Location/Qualifiers	
FEATURES			
source		1..859	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6196657"	
		/sex="male"	
		/tissue_type="sympathetic trunk"	
		/dev_stage="adult, 16 yr"	
		/lab_host="DH10B"	
		/clone_lib="Lupski syphathetic trunk"	
		/note="Vector: pCMV-SPORT5 (Life Technologies); Site_1: NotI; Site2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCG-3' and 5'-GATAGTTCTAGATCGGACGCGGCCCTT(15)-3'. Size selected 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."	
ORIGIN			
Query Match		18.4%; Score 822.8; DB 5; Length 859;	
Best Local Similarity		98.3%; Pred. No. 7.6e-177;	
Matches 843; Conservative		0; Mismatches 12; Indels 3; Gaps	
QY	1823	CTTGCCCATCTCTCTCTCGGCCACCTTCTCGCTACCAAGTCCCAACTCTCTTC	1882
Db	1	CTTGCCCATCTCTCTCTCGGCCACCTTCTCGCTACCAAGTCCCAACTCTCTTC	60
QY	1883	CACCTCGGGCAGCCACCCCTACAGCGCGGCCCTCCAGGCCCTTCGCGGCATCCAGAC	1942
Db	61	CACCTCGGGCAGCCACCCCTACAGCGCGGCCCTCCAGGCCCTTCGCGGCATCCAGAC	120
QY	1943	CCCGCTCTACCCCTCTGCTCATGGTTCTCCAGCTCTCTTACTATGCCTAAACGCA	2002

Db	841	TTTTACTGACTCTGCTATCAGTGCCTAAAGTGAATGGCGGACCCCAAGAGGACCTGG	900
Qy	1274	A--GCCCTGGGATCAGGTGAA---CTCACAGCCAATGAGAACTTGGAGCTTTGGAAAA	1328
Db	901	NAGCCCTGGGATCAGGGGAACCTCCAGGCCAAATAGGGAACCTTGAAGCTTTTGNA	960
RESULT 8			
LOCUS	BQ716397		
DEFINITION	AGENCOURT_8477984_Lupski_sympathetic_trunk_Homo_sapiens_cDNA_clone		
ACCESSION	IMAGE:619657.5	mRNA	linear
VERSION	BQ716397		
KEYWORDS	BQ716397.1	GI:21855294	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	1 (bases 1 to 859)		
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-romail.nih.gov CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13605 row: h column: 02 High quality sequence stop: 659.		
FEATURES			
source	1..859		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:619657"		
	/sex="male"		
	/tissue_type="sympathetic trunk"		
	/dev_stage="adult, 16 yr"		
	/lab_host="DH10B"		
	/clone_lib="Lupski_sympathetic_trunk"		
	/note="Vector: pCMV-Sport6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCAACGCGTCCG-3' and 5'-GACTAGTTCAGATCGAGCGCGCCCTT(15)-3'. Size selected 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."		
ORIGIN			
Query Match	18.4%;	Score 822.8;	DB 5;
Best Local Similarity	98.3%;	Pred. No. 7.6e-177;	Length 859;
Matches	843;	Conservative 0;	Mismatches 12; Indels 3; Gaps 1;
Qy	1823	CTTGCCCATCTCTCTTCTCTCGCCCACTTCTCGCTACGAGTCAGGTCCTTCTTC	1888
Db	1	CTTGCCCATCTCTCTTCTCTCGCCCACTTCTCGCTACGAGTCAGTCCCACTCTTC	60
Qy	1883	CACCTCGGCAGCCACCCCTACACGGCCGCCCTCCAGGCCCCCTTCGGCGCATCCAGAC	1942
Db	61	CACCTCGGCAGCCACCCCTACACGGCCGCCCTCCAGGCCCCCTTCGGCGCATCCAGAC	120
Qy	1943	CCCGCTCACCCCTCTGCTCATGGTTCTCAGCTCCGTCTCTACTATGCTTAACCGCA	2002

Db 121 CCCGCTCACCCCTCTGCTCATGTTCTCCAGTCTCCTGTTCTCTACTATGCTTAACGCA 180  
QY 2003 TGTCTTCAAGAGGCGCTCAAGGATGTGCCAAAGGCCCGCAGCAGCATCTCTGAAATCACA 2062  
Db 181 TGTCTTCAAGAGGCGCTCAAGGATGTGCCAAAGGCCCGCAGCAGCATCTCTGAAATCACA 240  
QY 2063 GAGTTTCTGCTGGAGGGTTCATATCCAGTGGCTAGAAATTTGTATCCCAACACCCAC 2122  
Db 241 GAGTTTCTGCTGGAGGGTTCATATCCAGTGGCTAGAAATTTGTATCCCAACACCCAC 300  
QY 2123 CCAGTGAAGCAGCTACTCTCCAGTAGCAAGACAGTCCCTCGGGGGGAAGCTGGTTCAT 2182  
Db 301 CCAGTGAAGCAGCTACTCTCCAGTAGCAAGACAGTCCCTCGGGGGGAAGCTGGTTCAT 360  
QY 2183 CAGTGGTCACTGGGGTTCAGAGTATCCCTTAAACATCATAGACCCAGGCTCCCGAC 2242  
Db 361 CAGTGGTCACTGGGGTTCAGAGTATCCCTTAAACATCATAGACCCAGGCTCCCGAC 420  
QY 2243 AGAACAGTATTGGAATACCCCGCAGTGGCCAGTCTTCTGCTTCTCCCAAGCTGGTATTA 2302  
Db 421 AGAACAGTATTGGAATACCCCGCAGTGGCCAGTCTTCTGCTTCTCCCAAGCTGGTATTA 480  
QY 2303 TTCCAACTGAAGCTGTGCAATGCTTATTCAGCTGCATCTCCTACGCTGCTAGTCTCTG 2362  
Db 481 TTCCAACTGAAGCTGTGCAATGCTTATTCAGCTGCATCTCCTACGCTGCTAGTCTCTG 540  
QY 2363 CATCGAACAGAGCTGTTTACCCCTTCTAGTAGGCTTAAAGATTCACAGGCTTCAAGATCAGA 2422  
Db 541 CATCGAACAGAGCTGTTTACCCCTTCTAGTAGGCTTAAAGATTCACAGGCTTCAAGATCAGA 600  
QY 2423 GGCAGAACTCTCTCGAGGGAATAAGAAATATTTAAACCCCAATGAAACATCACCTAGCT 2482  
Db 601 GGCAGAACTCTCTCGAGGGAATAAGAAATATTTAAACCCCAATGAAACATCACCTAGCT 660  
QY 2483 TCTCAAAAGCTCAAAACAAAGGTATATCACCAAGTGTGTTCTGAACATAGAAACAGATTG 2542  
Db 661 TCTCAAAAGCTCAAAACAAAGGTATATCACCAAGTGTGTTCTGAACATAGAAACAGATTG 720  
QY 2543 ATGATTTAAAGAAATTTAAGAAATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTA 2602  
Db 721 ATGATTTAAAGAAATTTAAGAAATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTA 780  
QY 2603 TGGATCAACTACTAAACAAATAAGAGGAGAGGAGAGAAATCAAGAGA----TTTGATCAAG 2659  
Db 781 TGGATCAACTACTAAACAAATAAGAGGAGAGGAGAGAAATCCAGGAGATTTGGATCAAG 840  
QY 2660 ACAAATTTGAACCAAGTG 2677  
Db 841 ACCAAATTTGAACCAAGTG 858

RESULT 9  
LOCUS CA489164  
DEFINITION AGENCOURT\_10853363 MAPcL Homo sapiens cDNA clone IMAGE:6721472 5', mRNA linear EST 14-NOV-2002  
ACCESSION CA489164  
VERSION CA489164.1 GI:24951955  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 889)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-romail.nih.gov](mailto:cgabbs-romail.nih.gov)  
Tissue Procurement: Kristi A. Egland, Ira Pastan  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM14282 Row: k Column: 08  
High quality sequence stop: 683.  
Location/Qualifiers  
1. 889  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6721472"  
/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HMEL, LNCaP"  
/lab\_host="EMDH10B"  
/clone\_lib="MAPcL"  
/note="Vector: PCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Egland, James J. Vincent, Robert Strausberg, Bungkook Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Query Match 18.3%; Score 819.2; DB 6; Length 889;  
Best Local Similarity 98.2%; Pred. No. 5.1e-176;  
Matches 851; Conservative 0; Mismatches 9; Indels 7; Gaps 2;  
QY 2373 AGCTGTTACCCCTTCTAGTAGGCTTAAAGATTCAGGCTTCAAGATCAGAGCGAAGCTC 2432  
Db 1 AGCTGTTACCCCTTCTAGTAGGCTTAAAGATTCAGGCTTCAAGATCAGAGCGAAGCTC 60  
QY 2433 TCCTCGAGGGGAATAAGAAATATTTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGC 2492  
Db 61 TCCTCGAGGGGAATAAGAAATATTTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGC 120  
QY 2493 TGAACCAAAAGGTATATCACCAAGTGTGTTCTGAACATAGAAACATGATGATTGATTTAAA 2552  
Db 121 TGAACCAAAAGGTATATCACCAAGTGTGTTCTGAACATAGAAACATGATGATTGATTTAAA 180  
QY 2553 GAAATTTAAGAAATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTATGGATCAACT 2612  
Db 181 GAAATTTAAGAAATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTATGGATCAACT 240  
QY 2613 ACTAAACAAATAAGAGAGGAGAAATCAAGAGATTTGATCAAGACAAATTTGAACC 2672  
Db 241 ACTAAACAAATAAGAGAGGAGAAATCAAGAGATTTGATCAAGACAAATTTGAACC 300  
QY 2673 AAGTCTAAGGATTTCTTTTATTGAAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAA 2732  
Db 301 AAGTCTAAGGATTTCTTTTATTGAAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAA 360  
QY 2733 GCCGAATAGCCCGCAGCATTTCCCTTCAATCTTAGTAAACGAGGACCAAGAGGGGACC 2792  
Db 361 GCCGAATAGCCCGCAGCATTTCCCTTCAATCTTAGTAAACGAGGACCAAGAGGGGACC 420  
QY 2793 TGAGTCACTTTCCCAAGGGTTTCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGA 2852  
Db 421 TGAGTCACTTTCCCAAGGGTTTCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGA 480  
QY 2853 TAAGGAAGAGAAAGACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATTCCTCAATGC 2912  
Db 481 TAAGGAAGAGAAAGACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATTCCTCAATGC 540  
QY 2913 AAAGAGTTCAACCCAGCTTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACC 2972  
Db 541 AAAGAGTTCAACCCAGCTTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACC 600  
QY 2973 TCGGCCTCAACACAACTACCCATCTATGGTGGTTCATCAACAGCAACTCCAGTTTAA 3032  
Db 601 TCGGCCTCAACACAACTACCCATCTATGGTGGTTCATCAACAGCAACTCCAGTTTAA 660

```
QY 3033 TACTCAGCCTGTTGTTGTTGTCACCAAAATATGATGATATCCAGTCCAGTGAGGCCAGGCGT 3092
Db 661 TACTCAGCCTGTTGTTGTTGTCACCAAAATATGATGATATCCAGTCCAGTGAGGCCAGGCGT 720
QY 3093 GCAACCTTTATACCAATACCTATGAGCCCATGAGCCAGTGAATCAAGCCAAAGACATATAG 3152
Db 721 GCAACCTTTATATACCAATACCTATGAGCCCATGAGCCAGTGAATCAAGCCAAAGACATATAG 779
QY 3153 AGC-----AGTACCAAAATATGCCCCCAACAGAGGGGCAAGACCATCATCATCAGAGTGCCTAT 3206
Db 780 AGCAGTAAAGTACCAAAATATGCCCCCAACAGAGGGGCAAGACCATCATCATCAGAGTGCCTAT 839
QY 3207 GATGCACCCAGGCTCAGCAGCGGGGCC 3233
Db 840 GATGCACCCAGGCTCAGCAGCGGGGCC 866

RESULT 10
AUI24606
LOCUS
DEFINITION AUI24606 NT2RM4 Homo sapiens cDNA clone NT2RM4000268 5', mRNA
sequence.
ACCESSION AUI24606
VERSION AUI24606
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 846)
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J.,
Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T.,
Suzuki,Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM4000268"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RM4"
/notes="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"

ORIGIN
Query Match 17.8%; Score 797.6; DB 1; Length 846;
Best Local Similarity 98.08; Pred. No. 4.4e-171;
Matches 828; Conservative 0; Mismatches 15; Indels 2; Gaps 2;
QY 2387 CTAGTGAGGCTAAAGATTCCAGGCTTCAA-GATCAGAGGCGAGAACTCTCTCGAGGGAAT 2445
Db 3 CTAGTGGGGCTAAAGATTCCAGGCTTCAAAGGATCAGAGGCGAGAACTCTCTCGAGGGAAT 62
QY 2446 AAGAAAATATTAACCCATGAACATCAGTCTCTCAAAAGCTGAAACAAAGGT 2505
```

```
Db 63 AAAGAAAATATTAACCCATGAACATCAGTCTCTCAAAAGCTGAAACAAAGT 122
QY 2506 ATATCACCAGTTGTTTCTGAAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAA 2565
Db 123 ATATCACCAGTTGTTTCTGAAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAA 182
QY 2566 GATTTTAGGTTACAGCCAAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAT 2625
Db 183 GATTTTAGGTTACAGCCAAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAT 242
QY 2626 AGAGAGGGAGAAAATCAAGAGATTGATCAAGAACAAAATTTGAACCAAGTCTAAGGAT 2685
Db 243 AGAGAGGGAGAAAATCAAGAGATTGATCAAGAACAAAATTTGAACCAAGTCTAAGGAT 302
QY 2686 TCTTTTCATTTGAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGATAGCCCC 2745
Db 303 TCTTTTCATTTGAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGATAGCCCC 362
QY 2746 AGCATTTCCCTTCAATAGTAAACACGGAGCAAGAGGGGACCTGAGGTCACTTCC 2805
Db 363 AGCATTTCCCTTCAATAGTAAACACGGAGCAAGAGGGGACCTGAGGTCACTTCC 422
QY 2806 CAAGGGGTTTCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGATAAGGAAGAGAAG 2865
Db 423 CAAGGGGTTTCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGATAAGGAAGAGAAG 482
QY 2866 AAAGACGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGAGTTCAAC 2925
Db 483 AAAGACGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGAGTTCAAC 542
QY 2926 CCAGCTTCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACTCGGCTCAAGCA 2985
Db 543 CCAGCTTCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACTCGGCTCAAGCA 602
QY 2986 CAACCTAGCCCCATCTATGTGGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGCCTGT 3045
Db 603 CAACCTAGCCCCATCTATGTGGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGCCTGT 661
QY 3046 TGTTTTGCACCAAAATATGATGATATCAGTCCAGTGAGCCAGGCGTGCAACCTTTATAC 3105
Db 662 TGTTTTGCACCAAAATATGATGATATCAGTCCAGTGAGCCAGGCGTGCAACCTTTATAC 721
QY 3106 CCAATACCTATGAGCCCATGCGAGTGAATCAAGCCAAAGACATATAGAGAGTACCAAT 3165
Db 722 CCAATACCTATGAGCCCATGCGAGTGAATCAAGCCAAAGACATATAGAGAGTACCAAT 781
QY 3166 ATGCCCAACAGCGGCAAGACCATCATCAGATGCCATCATCAGATGCCATGATGACCCAGGCTCAGCA 3225
Db 782 ATGCCCAACAGCGGCAAGACCATCATCAGATGCCATCATCAGATGCCATGATGACCCAGGCTCAGCA 841
QY 3226 GCGGG 3230
Db 842 GCGGG 846

RESULT 11
AL554863/c
LOCUS
DEFINITION AL554863 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1087Y117 3-PRIME, mRNA sequence.
ACCESSION AL554863
VERSION AL554863.3
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1066)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31276673.
```





QY	3234	ACCGATTGCAGCCACCCACACAGCTTACTCCAGCCAATATGTTGCCTACAGTCTCTCAGCA	3293
DB	1	ACCGATTGCAGCCACCCACACAGCTTACTCCAGCCAATATGTTGCCTACAGTCTCTCAGCA	60
QY	3294	GTTTCCCAAAATCAGCCCTCTGTTGAGCATGTGCCACATTTACGTCCTCAGCATCCTCATGT	3353
DB	61	GTTTCCCAAAATCAGCCCTCTGTTGAGCATGTGCCACATTTACGTCCTCAGCATCCTCATGT	120
QY	3354	CTATAGTCCTGTAAATACAGGGTAATGCTAGAAATGATGGCCACCAACACACAGCCCGCAGCC	3413
DB	121	CTATAGTCCTGTAAATACAGGGTAATGCTAGAAATGATGGCCACCAACACACAGCCCGCAGCC	180
QY	3414	TGGTTTATGATATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGGCATTCGGATGTA	3473
DB	181	TGGTTTATGATATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGGCATTCGGATGTA	240
QY	3474	TGCATGTCCCAAAATACCATACAAAGGAGACAAGCCCTCTCTTTCTACTTTTGCCATTTC	3533
DB	241	TGCATGTCCCAAAATACCATACAAAGGAGACAAGCCCTCTCTTTCTACTTTTGCCATTTC	300
QY	3534	CACGGGCTCCCTTGCTCAGCAGTATGGCCACCTTAACGCTACCTTGACCCACACATACTCC	3593
DB	301	CACGGGCTCCCTTGCTCAGCAGTATGGCCACCTTAACGCTACCTTGACCCACACATACTCC	360
QY	3594	ACACCTTCAGCCTTCAGTACCCCACTGGGACAGCAGGAAAGCCAAACATGCTGGAAGTCA	3653
DB	361	ACACCTTCAGCCTTCAGTACCCCACTGGGACAGCAGGAAAGCCAAACATGCTGGAAGTCA	420
QY	3654	TCCGTGCAACCCAGTCTGTTTCAGCACCACTACGACACAGCGGCCCGAGGCTCTCCATCTGGC	3713
DB	421	TCCGTGCAACCCAGTCTGTTTCAGCACCACTACGACACAGCGGCCCGAGGCTCTCCATCTGGC	480
QY	3714	CAGTCCACAGCAGCAGTCAGCCATTTTACACGCGGGGCTTGGCGCAACTCCACCTCCAT	3773
DB	481	CAGTCCACAGCAGCAGTCAGCCATTTTACACGCGGGGCTTGGCGCAACTCCACCTCCAT	540
QY	3774	GACACTGCTCTCGAACGCGAGTGGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGT	3833
DB	541	GACACTGCTCTCGAACGCGAGTGGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGT	600
QY	3834	CTTTTACGATCCATCTTCTCTCAGCTTCAGCGCGGTATACCAACCCACCCACATGGCCCA	3893
DB	601	CTTTTACGATCCATCTTCTCTCAGCTTCAGCGCGGTATACCAACCCACCCACATGGCCCA	659
QY	3894	CGTACTCAGGCTCATGTACAGTCA-GGAATGGTTCTTCTTCATCTCCAACTGCCCATGGCC	3952
DB	660	CGTACTCAGGCTCATGTACAGTCA-CAGGAAATGGTTCTTCTTCATCCAACTGCCCATGGCC	719
QY	3953	CAATGATGCTAATGACGACAGCCACCGGGGTCCCGAGGCGCCCTCGCTCAAAATG	4012
DB	720	CAATGATGCTAATGACGACAGCCACCGGGGTCCCGAGGCGCCCTCGCTCAAAATG	778
QY	4013	CACCTACAGCCCATTCAGTCTCGAACACAGGCGCATTTCCCTCTATATGACGCAACCTTCAG	4072
DB	779	CACCTACAG-CCATTCAGTCTCGAACACAGGCGCATTTCCCTCTATATGACGCAACCTTCAG	836
QY	4073	TACAAGCCCAACCA	4088
DB	837	TACAAGCCCAACCA	852

RESULT 13	CD616048	830 bp	mRNA	linear	EST 12-JAN-2004
LOCUS	CD616048/c				
DEFINITION	50669370H1 FLP Homo sapiens	cdna, mRNA sequence.			
ACCESSION	CD616048				
VERSION	CD616048.1	GI:40264312			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				

QY 3998 CCCTCGCTCAAAAGTGCACCTAGACGCCATTCAGTCTTCGACACAGCGCATTTCCCTCTATA 4057  
 |||||  
 Db 111 CCCTCGCTCAAAAGTGCACCTAGACGCCATTCAGTCTTCGACACAGCGCATTTCCCTCTATA 52  
 |||||  
 QY 4058 TGAACGACCCCTTCAGTACAAAGCCCAACCAACAG 4092  
 |||||  
 Db 51 TGACGACCCCTTCAGTACAAAGCCCAACCAACAG 17  
 |||||

RESULT 14  
 C0886647  
 LOCUS  
 DEFINITION Boven\_14972 normal cattle brain Bos taurus cDNA clone  
 RZPDp1056D1525Q 5', mRNA sequence.  
 C0886647  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

EST.  
 Bos taurus (cow)  
 Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
 Pecora; Bovidae; Bovinae; Bos.

REFERENCE  
 AUTHORS Hennig S., Janitz M., Herwig R. and Williams J.  
 TITLE Generation, annotation, evolutionary analysis and database  
 integration of 14969 cattle EST clusters  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Hennig S

laboraty 123, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Innestr. 63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1612  
 Fax: +49 30 8413 1380  
 Email: hennigmolgen.mpg.de

The library was characterised by oligonucleotide fingerprinting  
 (ONFP) to reduce sequencing redundancy. According to the ONFP  
 procedure, clones that display the same hybridisation matrix with a  
 battery of 200 8mer oligonucleotides are grouped into clusters. One  
 clone per ONFP cluster was selected for sequencing. cDNA clones and  
 filters are distributed via Deutsches Ressourcenzentrum fuer  
 Genomforschung GmbH (<http://www.rzpd.de>).

PCR Primers  
 FORWARD: 5' CCCAGGCTTACATTTATGCTTCGGCTCG 3' (M13RSP) 5'-seq  
 BACKWARD: 5' GCTATTAGCCAGCTGGCAAGGGGATGTG 3' (M13FSP) 3'-seq  
 Seq primer: 5'-CCGCTCCGGAATTCGGGT-3' (M13RSP).

FEATURES  
 Location/Qualifiers  
 1..877  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /clone="RZPDp1056D1525Q"  
 /sex="female"  
 /tissue\_type="brain tissue"  
 /dev\_stage="adult brain"  
 /clone\_lib="normal cattle brain"  
 /note="Organ: brain; Vector: pSport1; Site 1: NotI;  
 Site 2: SalI; Random primed and directionally cloned in  
 pSport1 vector using NotI  
 5'-PGACTAGTTCTAGATCGCGAGCGCGCC (T)15-3' and SalI 5'-  
 TCGACCCAGCGTCCG-3' adapters (Gibco BRL)";

ORIGIN  
 Query Match 17.0%; Score 761.2; DB 7; Length 877;  
 Best Local Similarity 95.6%; Pred. No. 9.5e-163;  
 Matches 805; Conservative 0; Mismatches 33; Indels 4; Gaps 2;

QY 3643 GGTGGAAGTCATCTCGACCCAGTCTCTGTTCAGCACCATCAGACAGCGCCCGCAGGCT 3702  
 |||||  
 Db 34 GGTGGAGGCATCTCTCGCCAGTCTCTGTTCAGCACCATCAGACAGCGCCCGCTCAGGCT 93  
 |||||  
 QY 3703 CTCCTCTCGCAGTCCACAGCAGCAGTCAGCCATTTACACGCGGGCTTCGGCCACT 3762  
 |||||  
 Db 94 CTCCTCTCGCAGTCCACAGCAGCAGTCAGCCATTTACCATGCGGGCTTCGGCCACC 153  
 |||||

QY 3763 CCACCTCCATGACACACCTGCTCCAAACACGCGAGTCGCACAGAAATAGTTTCCACGACGA 3822  
 |||||  
 Db 154 CCACCTCCATGACACACCTGCTCCAAATACACAGTCTCCACAGAAATAGTTTCCACGACGA 213  
 |||||  
 QY 3823 CAACAGACTGTCTTTAGATCCATCCTTCTACGCTTCAGCGCGGTATACCAACCCACCC 3882  
 |||||  
 Db 214 CAGCAGACCGTCTTTACCATTCATCCTTCTCATGTTTCAGCCGGCTTACACCAATCCACCC 273  
 |||||  
 QY 3883 CACATGGCCACGTCACCTACGCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCAACT 3942  
 |||||  
 Db 274 CACATGGCCATGTATACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCAACT 333  
 |||||  
 QY 3943 GCCATGCGCCAATGATCTAATGACGACACAGCCACCCGCGGTCCCGAGCGCCCTC 4002  
 |||||  
 Db 334 GCCCATGCGCCAATGATCTAATGACGACACAGCCACCCGCGGTCCCGAGCGCCCTC 393  
 |||||  
 QY 4003 GCTCAAGTGCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTCTATATGAG 4062  
 |||||  
 Db 394 GCTCAAGTGCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTCTATATGAG 453  
 |||||  
 QY 4063 CACCTTCAGTACAAAGCCACCAACACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAA 4122  
 |||||  
 Db 454 CACCTTCAGTACAAAGCCACCAACACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAA 513  
 |||||  
 QY 4123 AGGCCAAATCCCTCTCCTCTTCTACTGCTTCTACCAACTGGAAGCAGAGAAACTAGAA 4182  
 |||||  
 Db 514 AGGCCAAATCCCTCTCCTCTTCTACTGCTTCTACCAACTGGAAGCAGAGAAACTAGAA 573  
 |||||  
 QY 4183 TTTCAATTTATTTGTTTTTAAATATATATGTTGATTTCTTGTAAACATCCAATAGGAATG 4242  
 |||||  
 Db 574 TTTCA-TTATTTTGTGTTTTGAAAAAATATATGTTGATTTCTTGTAAACATCCAATAGGAATG 632  
 |||||  
 QY 4243 CTAACAGTTCATCTGCAAGGAGATACCTGACCGAGTAGAGGATTTAGGAACCTTGGG 4302  
 |||||  
 Db 633 CTAACAGTTCATCTGCAAGGAGATATTGACCGAGTAGAGGATTTAGGAACCTTGGG 692  
 |||||  
 QY 4303 GGCTATTCCTAATTCATATGCTGTTTCAAGTCCCGCAGGTACCCAGCTCTGCTTGC 4362  
 |||||  
 Db 693 GGCTGTTCCATAATTCATATGCTGTTTCCAAGTCCCGCAGGTACCCAGCTCTGCTTGC 752  
 |||||  
 QY 4363 CGAACTGGAAGTATTTATTTTAAATTAACCTTGAAGTCATG---AACACATCACT 4419  
 |||||  
 Db 753 TGAACACTGGAAGTATTTATTTATTTTATGACCTTTGAAGATCATGAACAACATCACT 812  
 |||||  
 QY 4420 AGCAAAAGAGTAACAAGAGTATTTCTTGTGCTATTTACTGCTAAATAAAAAAAAAA 4479  
 |||||  
 Db 813 AGCAAAAGAGTAACAAGAGTATTTCTTGTCTATTACCGCTAAAAAAAAAAAAAAAA 872  
 |||||  
 QY 4480 AA 4481  
 |||||  
 Db 873 AA 874

RESULT 15  
 CA775370/c  
 LOCUS  
 DEFINITION io88f10.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6133411 3',  
 similar to TR:Q99700 Q99700 ATAXIN-2. ;, mRNA sequence.  
 ACCESSION CA775370.1 GI:26012904  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 784)  
 REFERENCE  
 AUTHORS Melton D., Brown J., Kent G., Permutt A., Lee C., Kaestner K.,  
 Lemishka I., Scarce M., Brestelli J., Gradwohl G., Clifton S.,  
 Hillier L., Marra M., Pape D., Wylie T., Martin J., Blaisain A.,  
 Schmitt A., Theising B., Ritter E., Ronko I., Bennett J.,  
 Cardenas M., Gibbons M., McCann R., Cole R., Tsagarisshvili R.,



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 17:59:31 ; Search time 776 Seconds  
(without alignments)  
10264.502 Million cell updates/sec

Title: US-10-802-228-1  
Perfect score: 4481  
Sequence: 1 acccccgagaaagcaaccacccca.....taaaaaaaaaaaaaaaaaaaaaa 4481

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: +  
1: /cgn2\_6/prodata/1/ina/1\_COMB.seq: +  
2: /cgn2\_6/prodata/1/ina/5\_COMB.seq: +  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq: +  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq: +  
5: /cgn2\_6/prodata/1/ina/H\_COMB.seq: +  
6: /cgn2\_6/prodata/1/ina/PTUS\_COMB.seq: +  
7: /cgn2\_6/prodata/1/ina/PP\_COMB.seq: +  
8: /cgn2\_6/prodata/1/ina/RE\_COMB.seq: +  
9: /cgn2\_6/prodata/1/ina/backfiles1.seq: +

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4481	100.0	4481	3	US-09-041-886-18
2	4481	100.0	4481	3	US-09-648-281-1
3	4481	100.0	4481	3	US-09-707-919A-20
4	4481	100.0	4481	3	US-09-083-268-2
5	4481	100.0	4481	3	US-08-981-998A-2
6	4447	99.2	4484	3	US-09-949-016-4498
7	2957	66.0	3798	3	US-08-981-998A-4
8	935.2	20.9	1257	3	US-09-648-281-11
9	935.2	20.9	1257	3	US-09-083-268-4
10	935.2	20.9	1257	3	US-08-981-998A-19
11	870	19.4	151088	3	US-09-949-016-16240
12	571.2	12.7	623	3	US-09-043-303-5
13	497.4	11.1	516	3	US-09-083-268-1
14	497.4	11.1	516	3	US-08-981-998A-1
15	455.8	10.2	459	3	US-09-707-919A-13
16	412.4	9.2	49677	3	US-09-949-002-817
17	354.2	7.9	355	3	US-09-043-303-1
18	313.2	7.0	47677	3	US-09-949-002-668
19	234	5.2	601	3	US-09-949-016-160682
20	183	4.1	183	3	US-09-707-919A-16
21	177	4.0	177	3	US-09-707-919A-17
22	169	3.8	169	3	US-09-707-919A-15
23	167	3.7	203	3	US-09-043-303-3
24	149	3.3	601	3	US-09-949-016-160680

C	25	149	3.3	601	3	US-09-949-016-160681	Sequence 160681,
	26	106	2.4	106	3	US-09-707-919A-14	Sequence 14, Appl
	27	99.4	2.2	1614	3	US-09-616-289-45	Sequence 45, Appl
	28	99.4	2.2	1614	3	US-09-976-740-45	Sequence 45, Appl
	29	99.4	2.2	12425	3	US-09-616-289-50	Sequence 50, Appl
	30	99.4	2.2	12425	3	US-09-976-740-50	Sequence 50, Appl
	31	98.8	2.2	2561	3	US-09-616-289-48	Sequence 48, Appl
	32	98.8	2.2	2561	3	US-09-976-740-48	Sequence 48, Appl
	33	98	2.2	98	3	US-09-707-919A-18	Sequence 18, Appl
	34	94.4	2.1	15695	3	US-09-949-016-15644	Sequence 15644, A
	35	93	2.1	3169	3	US-09-620-312D-713	Sequence 713, App
	36	87.8	2.0	477	3	US-09-135-994-1	Sequence 1, Appli
	37	87.8	2.0	477	3	US-09-684-843A-1	Sequence 1, Appli
	38	87	1.9	601	3	US-09-949-016-160684	Sequence 160684,
	39	86.2	1.9	8438	2	US-07-945-283-1	Sequence 1, Appli
	40	86.2	1.9	37875	3	US-09-949-016-13182	Sequence 13182, A
	41	85	1.9	12001	2	US-08-458-568A-11	Sequence 11, Appl
	42	81.4	1.8	10348	2	US-08-457-273B-41	Sequence 41, Appl
	43	81.4	1.8	10348	3	US-08-556-419-13	Sequence 13, Appl
	44	81.4	1.8	10348	3	US-09-041-886-14	Sequence 14, Appl
	45	81.4	1.8	10366	2	US-08-246-982A-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-09-041-886-18  
; Sequence 18, Application US/09041886  
; Patent No. 6235872  
; GENERAL INFORMATION:  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Rabizadeh, Sharroz  
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence  
; TITLE OF INVENTION: Polypeptides and Methods of Use  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/041,886  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 2626  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4481 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 163..4099  
US-09-041-886-18  
Query Match 100.0%; Score 4481; DB 3; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ACCCCGAGAAAGCAACCCAGCGCCGCCCGCTCTCTACGTTGTCCTTCCGCGCCCGGG	60
Db	1	ACCCCGAGAAAGCAACCCAGCGCCGCCCGCTCTCTACGTTGTCCTTCCGCGCCCGGG	60
QY	61	GCCACCTCAGCTTCGCTTCGGTCTGACGCCCTCCGACTTCGGTAAAGTAGTCCCTATCCG	120
Db	61	GCCACCTCAGCTTCGCTTCGGTCTGACGCCCTCCGACTTCGGTAAAGTAGTCCCTATCCG	120
QY	121	CACCTCCGCTCCACCCCGCGCTCGGCGCGCCCTCCGATCGCTCGATCGCTCAGCGSCGCA	180
Db	121	CACCTCCGCTCCACCCCGCGCTCGGCGCGCCCTCCGATCGCTCGATCGCTCAGCGSCGCA	180
QY	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTTCGCGCTTCGCGCAGCCAGGTGGGCC	240
Db	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTTCGCGCTTCGCGCGCAGCCAGGTGGGCC	240
QY	241	GGGTGGCGCTCGCTCAGGGCGCGCGCGGAGCGGGCGGGCGGGCGGTGGCGCGCC	300
Db	241	GGGTGGCGCTCGCTCAGGGCGCGCGCGGAGCGGGCGGGCGGGCGGTGGCGCGCC	300
QY	301	CGGGACCGTATCCCTCCGCGCCCTCCCGCGCGCGCCCGCGCGCCCTCCCTCCCGG	360
Db	301	CGGGACCGTATCCCTCCGCGCCCTCCCGCGCGCGCCCGCGCGCCCTCCCTCCCGG	360
QY	361	CAGAGCTCGCCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCGG	420
Db	361	CAGAGCTCGCCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCGG	420
QY	421	TTTGGGCGGCTCCGGGGCTCTTGGTCTCGCGGGCGCTCCCGCGCCCTTCGTGTC	480
Db	421	TTTGGGCGGCTCCGGGGCTCTTGGTCTCGCGGGCGCTCCCGCGCCCTTCGTGTC	480
QY	481	GTCTTCTCCCTCGCAGCCCGGGCGCCCTCCGCGCGCGCCCAACCGCGCTCCCGG	540
Db	481	GTCTTCTCCCTCGCAGCCCGGGCGCCCTCCGCGCGCGCCCAACCGCGCTCCCGG	540
QY	541	CTGGCGCCGTGCTCCCGCGGTTCGGCGTCTCTTGGCGCGCCCGCTCCCGGC	600
Db	541	CTGGCGCCGTGCTCCCGCGGTTCGGCGTCTCTTGGCGCGCCCGCTCCCGGC	600
QY	601	TGTCCCGCGCGCTCGAGCGGTGTATGGGCCCTTCAACATGCTGAGCGCCCGAG	660
Db	601	TGTCCCGCGCGCTCGAGCGGTGTATGGGCCCTTCAACATGCTGAGCGCCCGAG	660
QY	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG	720
Db	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG	720
QY	721	CAGCGCGCGCGGCTGCCAATCGCGCAAGCCGGCGAGCGGCTTCTAGCGTCCG	780
Db	721	CAGCGCGCGCGGCTGCCAATCGCGCAAGCCGGCGAGCGGCTTCTAGCGTCCG	780
QY	781	CCGCGCGCGCGCTTCGCGCTCTGCTCGTCTCGTCTCTCGGCGCGGCTCC	840
Db	781	CCGCGCGCGCGCTTCGCGCTCTGCTCGTCTCTCGGCGCGGCTCC	840
QY	841	TCCTCGGTGTCGCGCGACCTTCGCGCGCGGGAGCGCGCTGGGCGAGGTCGAAC	900
Db	841	TCCTCGGTGTCGCGCGACCTTCGCGCGCGGGAGCGCGCTGGGCGAGGTCGAAC	900
QY	901	AGTAAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGTAGAAATCTATGCAAAATAGG	960
Db	901	AGTAAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGTAGAAATCTATGCAAAATAGG	960
QY	961	ATGGTTCAATATCTACATCAGTTGTCGCTCAAAATGTGAAGTACAAAGTGA	1020
Db	961	ATGGTTCAATATCTACATCAGTTGTCGCTCAAAATGTGAAGTACAAAGTGA	1020
QY	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGCTATGATGC	1080
Db	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGCTATGATGC	1080

QY	1081	GCACATGAGAAAGGTACAGAAATCCAGTTCGGGGCGAAACGTGAAGAAATATAGGAGT	1140
Db	1081	GCACATGAGAAAGGTACAGAAATCCAGTTCGGGGCGAAACGTGAAGAAATATAGGAGT	1140
QY	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTGTACAGTTTAAAGATATGGACTCCAGTTAT	1200
Db	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTGTACAGTTTAAAGATATGGACTCCAGTTAT	1200
QY	1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGCGAACACAAA	1260
Db	1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGCGAACACAAA	1260
QY	1261	GAGAAGGACCTGGAGCCCTGGGATGCAAGTGAATCAGAGCAATAGAGAACTTTGAGCT	1320
Db	1261	GAGAAGGACCTGGAGCCCTGGGATGCAAGTGAATCAGAGCAATAGAGAACTTTGAGCT	1320
QY	1321	TTGGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
Db	1321	TTGGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
QY	1381	AATTATGTTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA	1440
Db	1381	AATTATGTTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA	1440
QY	1441	AGAGATAACTCAGAAAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA	1500
Db	1441	AGAGATAACTCAGAAAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA	1500
QY	1501	ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTTGGAAAAATGATAGGAGTGAG	1560
Db	1501	ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTTGGAAAAATGATAGGAGTGAG	1560
QY	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGCACAGATAAAC	1620
Db	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGCACAGATAAAC	1620
QY	1621	ACTAGGGAAAAATAATATATTCCTCTCGGACAAAGAAATAGAGAAATCATATCTCTGGGA	1680
Db	1621	ACTAGGGAAAAATAATATATTCCTCTCGGACAAAGAAATAGAGAAATCATATCTCTGGGA	1680
QY	1681	AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
Db	1681	AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
QY	1741	TCCACTTCTCACACTTCAGATTTTCAACCCGAAATTCGGTTCAGACCCAAAGAGTAGTTAAT	1800
Db	1741	TCCACTTCTCACACTTCAGATTTTCAACCCGAAATTCGGTTCAGACCCAAAGAGTAGTTAAT	1800
QY	1801	GGAGTGTTCCTCGGCCATCGCTTGGCCCATCTCTCTCGGCCACCTTCTCGCTAC	1860
Db	1801	GGAGTGTTCCTCGGCCATCGCTTGGCCCATCTCTCTCGGCCACCTTCTCGCTAC	1860
QY	1861	CAGTCAGTCCCAACTCTTTCACCTCGGGCAGCACCCCTACACGGCCGCGCTCCAGG	1920
Db	1861	CAGTCAGTCCCAACTCTTTCACCTCGGGCAGCACCCCTACACGGCCGCGCTCCAGG	1920
QY	1921	CCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGTCTCT	1980
Db	1921	CCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGTCTCT	1980
QY	1981	GTCTCTACTATGTCCTAAACGCATGTTTCAGAAAGGGCTCCCAAGATGTCTCCCAAAGGCC	2040
Db	1981	GTCTCTACTATGTCCTAAACGCATGTTTCAGAAAGGGCTCCCAAGATGTCTCCCAAAGGCC	2040
QY	2041	CAGGCACATCTCGAAATCAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCCTA	2100
Db	2041	CAGGCACATCTCGAAATCAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCCTA	2100
QY	2101	GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCACT	2160
Db	2101	GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCACT	2160

QY 2161 CCTCGGGGGAAAGTGGTTCATAGTGGTTCAGTGGGTTCCAGATTATCCCTAAAAC 2220  
DB 2161 CCTCGGGGGAAAGTGGTTCATAGTGGTTCAGTGGGTTCCAGATTATCCCTAAAAC 2220  
QY 2221 CATAGACCAGGTCTCCAGACAGAACAGTATGGAAATPACCCAGTGGGCGAGTTCTT 2280  
DB 2221 CATAGACCAGGTCTCCAGACAGAACAGTATGGAAATPACCCAGTGGGCGAGTTCTT 2280  
QY 2281 GCTTCTCCCAAGCTGGTATTATTCCACTGAAGCTGTGGCCATGCTATTCCAGCTGCA 2340  
DB 2281 GCTTCTCCCAAGCTGGTATTATTCCACTGAAGCTGTGGCCATGCTATTCCAGCTGCA 2340  
QY 2341 TCTCTACGCCCTGTAGTCTCGATCGAACAGAGCTGTTACCCCTTCTAGTCAGGCTAAA 2400  
DB 2341 TCTCTACGCCCTGTAGTCTCGATCGAACAGAGCTGTTACCCCTTCTAGTCAGGCTAAA 2400  
QY 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGACAGGAAATAAGAAATATTTAA 2460  
DB 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGACAGGAAATAAGAAATATTTAA 2460  
QY 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAAAGAGTTATACACAGTTGTT 2520  
DB 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAAAGAGTTATACACAGTTGTT 2520  
QY 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGNATGATTTTAGTTTACAG 2580  
DB 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGNATGATTTTAGTTTACAG 2580  
QY 2581 CCAAGTTCTACTTCTCAATCTATGATCAACTACTATAACAAATAAGAGGAGAGAAAA 2640  
DB 2581 CCAAGTTCTACTTCTCAATCTATGATCAACTACTATAACAAATAAGAGGAGAGAAAA 2640  
QY 2641 TCAAGAGATTGTATCAAGACAAAATTGAACCAAGTCTTCAAGGATTTCTTTTCAATTGAAAT 2700  
DB 2641 TCAAGAGATTGTATCAAGACAAAATTGAACCAAGTCTTCAAGGATTTCTTTTCAATTGAAAT 2700  
QY 2701 AGCAGAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTTCA 2760  
DB 2701 AGCAGAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTTCA 2760  
QY 2761 ATACTTGTAAACAGGAGCAAGAGGGACCTGAGTCACTTCCCAAGGGTTTCAAGCT 2820  
DB 2761 ATACTTGTAAACAGGAGCAAGAGGGACCTGAGTCACTTCCCAAGGGTTTCAAGCT 2820  
QY 2821 TCCAGGCCAGCATGTATAAAGAGAAAGACGATAGAGAGAGAGAAAGACGAGCTGAG 2880  
DB 2821 TCCAGGCCAGCATGTATAAAGAGAAAGACGATAGAGAGAGAGAAAGACGAGCTGAG 2880  
QY 2881 CAAGTTAGGAATCAACATTGAATCCCAATGCAAGAGAGTTCAACCAAGTTCTCTCT 2940  
DB 2881 CAAGTTAGGAATCAACATTGAATCCCAATGCAAGAGAGTTCAACCAAGTTCTCTCT 2940  
QY 2941 CAGCCAAAGCTTCTACTACCCCACTTCACTCCGCTTCAAGCACAACCTAGCCCATCT 3000  
DB 2941 CAGCCAAAGCTTCTACTACCCCACTTCACTCCGCTTCAAGCACAACCTAGCCCATCT 3000  
QY 3001 ATGTTGGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGCTCTGTTTGTGACCAAAAT 3060  
DB 3001 ATGTTGGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGCTCTGTTTGTGACCAAAAT 3060  
QY 3061 ATGATGTATCCAGTCCAGTGAGCCGAGGCTGGAACCTTTATACCAATACCTATGACG 3120  
DB 3061 ATGATGTATCCAGTCCAGTGAGCCGAGGCTGGAACCTTTATACCAATACCTATGACG 3120  
QY 3121 CCCATGCGAGTGAATCAAGCCACAGACATATAGAGCAGTACCAATATGCCCCACAGCGG 3180  
DB 3121 CCCATGCGAGTGAATCAAGCCACAGACATATAGAGCAGTACCAATATGCCCCACAGCGG 3180  
QY 3181 CAAGACAGCATCATCAGAGTGCCATGATGACCCAGCGTCAGCAGGGGCCACCGATT 3240  
DB 3181 CAAGACAGCATCATCAGAGTGCCATGATGACCCAGCGTCAGCAGGGGCCACCGATT 3240  
QY 3241 GCAGGCCACCCACAGCTTACTCCACGCAATATGTTGCTTACAGTCTCAGCAGTTCCCA 3300

DB 3241 GCAGGCCACCCACAGCTTACTCCAGCAATATGTTGCTACAGTCTCAGCAGTTCCCA 3300  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCAATTATCAGTCTCAGCATCTCTCATGTCTATAGT 3360  
DB 3301 AATCAGCCCTTGTTCAGCATGTGCCAATTATCAGTCTCAGCATCTCTCATGTCTATAGT 3360  
QY 3361 CTTGTAAATACAGGGTAAATGCTAGATGATGACCAACCAACAGCCAGCCTGGTTTA 3420  
DB 3361 CTTGTAAATACAGGGTAAATGCTAGATGATGACCAACCAACAGCCAGCCTGGTTTA 3420  
QY 3421 GTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACGATGCGATGTATGATGT 3480  
DB 3421 GTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACGATGCGATGTATGATGT 3480  
QY 3481 CCCAAATTAACATACCAAGAGAGCAAGCCCTTCTTCTACTTTTGCCATTTCCACGGGC 3540  
DB 3481 CCCAAATTAACATACCAAGAGAGCAAGCCCTTCTTCTACTTTTGCCATTTCCACGGGC 3540  
QY 3541 TCCCTTGTCTCAGCATGTGCCACCTTAACGCTACCTTGCACCCACATACTCCACACCT 3600  
DB 3541 TCCCTTGTCTCAGCATGTGCCACCTTAACGCTACCTTGCACCCACATACTCCACACCT 3600  
QY 3601 CAGCCTTCAAGCTTACCCCACTGACAGCAGCAAAAGCAACATGGTGGAAAGTCTCTCTCA 3660  
DB 3601 CAGCCTTCAAGCTTACCCCACTGACAGCAGCAAAAGCAACATGGTGGAAAGTCTCTCTCA 3660  
QY 3661 CCCAGTCTCTTTCAGCACCATCAGACCCAGGCGGCCAGGCTCTCCATCTGCCAGTCCA 3720  
DB 3661 CCCAGTCTCTTTCAGCACCATCAGACCCAGGCGGCCAGGCTCTCCATCTGCCAGTCCA 3720  
QY 3721 CAGCAGCAGTACGACCATTTACACGGGGGCTTGGCCAACTCCACCTTCCATGACACCT 3780  
DB 3721 CAGCAGCAGTACGACCATTTACACGGGGGCTTGGCCAACTCCACCTTCCATGACACCT 3780  
QY 3781 GCCTCCAAACAGCCAGTGCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
DB 3781 GCCTCCAAACAGCCAGTGCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840  
QY 3841 ATCCATCTCTTCTCAGTTCAGCGGGGTATACCAACCCACCCACATGCCCCACGTACCT 3900  
DB 3841 ATCCATCTCTTCTCAGTTCAGCGGGGTATACCAACCCACCCACATGCCCCACGTACCT 3900  
QY 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTTCTCATCCTCAACTGCCCATGGCCAAATGATG 3960  
DB 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTTCTCATCCTCAACTGCCCATGGCCAAATGATG 3960  
QY 3961 CTAATGACGACACAGCCCGGCTCCCGAGGCGCCCTCGCTCAAAAGTGCACCTACAG 4020  
DB 3961 CTAATGACGACACAGCCCGGCTCCCGAGGCGCCCTCGCTCAAAAGTGCACCTACAG 4020  
QY 4021 CCCATTCCAGTCTCGAACAACAGCGCATTTCCCTTATATGACGCAACCTTTCAGTACAAGCC 4080  
DB 4021 CCCATTCCAGTCTCGAACAACAGCGCATTTCCCTTATATGACGCAACCTTTCAGTACAAGCC 4080  
QY 4081 CACCACCAACAGCAGTTGTAAGGCTCCCTTGAGGAAACCGAAAGGCCAAATTCCTCTCTC 4140  
DB 4081 CACCACCAACAGCAGTTGTAAGGCTCCCTTGAGGAAACCGAAAGGCCAAATTCCTCTCTC 4140  
QY 4141 CCTTCTACTGCTTCTACCAACTGGAAGACAGAAAACTAGAAATTTTCAATTTATTTGTTTT 4200  
DB 4141 CCTTCTACTGCTTCTACCAACTGGAAGACAGAAAACTAGAAATTTTCAATTTATTTGTTTT 4200  
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGCTAAACAGTTCATTTGCAG 4260  
DB 4201 TAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGCTAAACAGTTCATTTGCAG 4260  
QY 4261 TGAAGATATCTTGGACCGAGTAGAGCATTTAGGAACCTTTGGGGGCTTATTTCCATTAATTTCCA 4320  
DB 4261 TGAAGATATCTTGGACCGAGTAGAGCATTTAGGAACCTTTGGGGGCTTATTTCCATTAATTTCCA 4320  
QY 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGGCCGAAACTGGAAGTATTTT 4380



```
Db 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGC CGAAA CTGGAAGTTATTT 4380
Qy 4381 ATTTTTTAATACCCCTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAACAAGAGT 4440
Db 4381 ATTTTTTAATACCCCTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAACAAGAGT 4440
Qy 4441 GATTCTTGCTGCTATCTACTGCTTAAATAAAAAAAAAAAAAA 4481
Db 4441 GATTCTTGCTGCTATCTACTGCTTAAATAAAAAAAAAAAAAA 4481

RESULT 2
US-09-648-281-1
; Sequence 1, Application US/09648281
; Patent No. 6515197
; GENERAL INFORMATION:
; APPLICANT: Pulet, Stefan M.
; TITLE OF INVENTION: Transgenic Animal Model of
; FILE OF INVENTION: Neurodegenerative Disease and Methods of Use
; FILE REFERENCE: P-CE 4336
; CURRENT APPLICATION NUMBER: US/09/648,281
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)...(4099)
US-09-648-281-1

Query Match 100.0%; Score 4481; DB 3; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCTCTCACGTGTCCTCCCGCGGCCCGGG 60
Db 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCCGCTCTCTCACGTGTCCTCCCGCGGCCCGGG 60

Qy 61 GCCACCTCAGCTTCGCTTCGCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
Db 61 GCCACCTCAGCTTCGCTTCGCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120

Qy 121 CACCTCCGCTCCACCCGCGGCTCGCGCGCCGCCCTCCGATCGCTCAGCGGCCGCA 180
Db 121 CACCTCCGCTCCACCCGCGGCTCGCGCGCCGCCCTCCGATCGCTCAGCGGCCGCA 180

Qy 181 GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTGGCCGACCGAGGTGGCCC 240
Db 181 GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTGGCCGACCGAGGTGGCCC 240

Qy 241 GGGTGGCGCTCGCTCCAGGGCGCGGCGGAGCGGCGGCGGCGGCGGTGGCGGCC 300
Db 241 GGGTGGCGCTCGCTCCAGGGCGCGGCGGAGCGGCGGCGGCGGCGGTGGCGGCC 300

Qy 301 CCGGACCGTATCCCTCCCGCGCCCTCCCGCGCCCGCGGCCCTCCCTCCCTCCCGG 360
Db 301 CCGGACCGTATCCCTCCCGCGCCCTCCCGCGCCCGCGGCCCTCCCTCCCTCCCGG 360

Qy 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTTGTAGCAACGGCAACGGCGGCGCGCG 420
Db 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTTGTAGCAACGGCAACGGCGGCGCGCG 420

Qy 421 TTTCCGCGCGCTCCGCGGCTCTTGGTCTCGGCGGCGCTCCCGCGCCCTTGGTCTC 480
Db 421 TTTCCGCGCGCTCCGCGGCTCTTGGTCTCGGCGGCGCTCCCGCGCCCTTGGTCTC 480

Qy 481 GTCCCTTCTCCCTCCGCGCGGCGCCCTCCCGCGCGCGCCCAACCGCGGCTCCCGG 540
Db 481 GTCCCTTCTCCCTCCGCGCGGCGCCCTCCCGCGCGCGCCCAACCGCGGCTCCCGG 540
```

```
Qy 541 CTCGGCGCCGCTGCGTCCCCCGCCGCTTCGCGCGCTCTCCCTTGGCGCGCCCGGCTCCCGGC 600
Db 541 CTCGGCGCCGCTGCGTCCCCCGCCGCTTCGCGCGCTCTCCCTTGGCGCGCCCGGCTCCCGGC 600

Qy 601 TGTCTCCCGCCGCGTGGAGCGGTGTATGGGCCCTTCCACCATGTCTGCTGAAGCCCGAG 660
Db 601 TGTCTCCCGCCGCGTGGAGCGGTGTATGGGCCCTTCCACCATGTCTGCTGAAGCCCGAG 660

Qy 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAG 720
Db 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAG 720

Qy 721 CAGCCGCGCCGCGCTGCGCAATGTCCGCAAGCCCGCGCGCGCGCGCTTCTTAGCGTTCG 780
Db 721 CAGCCGCGCCGCGCTGCGCAATGTCCGCAAGCCCGCGCGCGCGCTTCTTAGCGTTCG 780

Qy 781 CCCGCGCGCGCGCTTCCGCGTCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 840
Db 781 CCCGCGCGCGCGCTTCCGCGTCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 840

Qy 841 TCCTCGGTGTCGCGCGACCTCCGCGCGGAGGCGCGCGCTGGGCGAGAGTTCGAAAC 900
Db 841 TCCTCGGTGTCGCGCGACCTCCGCGCGGAGGCGCGCGCTGGGCGAGAGTTCGAAAC 900

Qy 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAG 960
Db 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAG 960

Qy 961 ATGGTTTATATATCTACTATCATCAGTTGTTGGCTCCAAATGTGAAGTACAAAGTACAA 1020
Db 961 ATGGTTTATATATCTACTATCATCAGTTGTTGGCTCCAAATGTGAAGTACAAAGTACAA 1020

Qy 1021 GGTATATATGAAGAGTGTAAAACTTACAGTCCGAAAGTGTGATTTGGTACTTGGATGCC 1080
Db 1021 GGTATATATGAAGAGTGTAAAACTTACAGTCCGAAAGTGTGATTTGGTACTTGGATGCC 1080

Qy 1081 GCAATGAGAAAGTACAGAAATCAGTTTCGGGCGGAAACGTGAAGAAATATGAGAGT 1140
Db 1081 GCAATGAGAAAGTACAGAAATCAGTTTCGGGCGGAAACGTGAAGAAATATGAGAGT 1140

Qy 1141 ATTTTGTTCAAATGTTTGTGTGTCAGTTTAAAGATATGGAATATGAGTCCAGTTAT 1200
Db 1141 ATTTTGTTCAAATGTTTGTGTGTCAGTTTAAAGATATGGAATATGAGTCCAGTTAT 1200

Qy 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAACACAAA 1260
Db 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAACACAAA 1260

Qy 1261 GAGAAAGACCTGGAGCCCTGGGATGCGAGTGAATCTCAGCCAATGAGGAACCTTGAGGCT 1320
Db 1261 GAGAAAGACCTGGAGCCCTGGGATGCGAGTGAATCTCAGCCAATGAGGAACCTTGAGGCT 1320

Qy 1321 TTGAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTTCGATATATGAAGAA 1380
Db 1321 TTGAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTTCGATATATGAAGAA 1380

Qy 1381 AATTATGCTGATGCTTACGTATGATAGAGTGTATCTTCTGATATCAGTCCCTTTAGAA 1440
Db 1381 AATTATGCTGATGCTTACGTATGATAGAGTGTATCTTCTGATATCAGTCCCTTTAGAA 1440

Qy 1441 AGAGATACTCAGAAAGATTTTTTAAACGGGAAGCAAGGGCAACACAGTTAGCAGAGAA 1500
Db 1441 AGAGATACTCAGAAAGATTTTTTAAACGGGAAGCAAGGGCAACACAGTTAGCAGAGAA 1500

Qy 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATAGGAGTCAG 1560
Db 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATAGGAGTCAG 1560

Qy 1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCGACAGCATAAAC 1620
Db 1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCGACAGCATAAAC 1620

Qy 1621 ACTAGGGAATAAATATATATATTCCTCGGACAAAGAAATAGAGAGTCAATCTCTGGGA 1680
```



3841 ATCCATCCTTCTCAGCTTCAGCGCGGTATACCAACCCACCCCAATGGCCCACTACT 3900  
Db  
3841 ATCCATCCTTCTCAGCTTCAGCGCGGTATACCAACCCACCCCAATGGCCCACTACT 3900  
Qy  
3901 CAGGCTCATGTACAGTACAGGATGTTCTTCTCATCCAACTGCCCATGGCCCATGATG 3960  
Db  
3901 CAGGCTCATGTACAGTACAGGATGTTCTTCTCATCCAACTGCCCATGGCCCATGATG 3960  
Qy  
3961 CTAATGAGACACAGCCACCCGCGGTCCCAAGGCGCCCTCGCTCAAAGTGCACTACAG 4020  
Db  
3961 CTAATGAGACACAGCCACCCGCGGTCCCAAGGCGCCCTCGCTCAAAGTGCACTACAG 4020  
Qy  
4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTATATAGACACCCCTTCAGTCAAGCC 4080  
Db  
4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTATATAGACACCCCTTCAGTCAAGCC 4080  
Qy  
4081 CACCAACCAACAGCAGTGTGAGGCTGCGCTGGAGGACCGAAGGCCAAATTCCTCCTC 4140  
Db  
4081 CACCAACCAACAGCAGTGTGAGGCTGCGCTGGAGGACCGAAGGCCAAATTCCTCCTC 4140  
Qy  
4141 CTTCTACTGCTTCTACCAACTGGAAGCACAGAAAAGTAAATTTTATTTTGTGTTT 4200  
Db  
4141 CTTCTACTGCTTCTACCAACTGGAAGCACAGAAAAGTAAATTTTATTTTGTGTTT 4200  
Qy  
4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTGCGAG 4260  
Db  
4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTGCGAG 4260  
Qy  
4261 TGAAGATACATTGGACCGAGTAGAGCATTTAGGAACCTTGGGGGTATTTCCATAATCCA 4320  
Db  
4261 TGAAGATACATTGGACCGAGTAGAGCATTTAGGAACCTTGGGGGTATTTCCATAATCCA 4320  
Qy  
4321 TATGCTGTTTACAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380  
Db  
4321 TATGCTGTTTACAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380  
Qy  
4381 ATTTTAAATACCTTGAAGTCAATGAAACATCAGCTAGCAAAAGAAAGTAAAGAGT 4440  
Db  
4381 ATTTTAAATACCTTGAAGTCAATGAAACATCAGCTAGCAAAAGAAAGTAAAGAGT 4440  
Qy  
4441 GATTTCTGCTGCTATTTACTGCTTAAATATATATATATATATATATATATATAT 4481  
Db  
4441 GATTTCTGCTGCTATTTACTGCTTAAATATATATATATATATATATATATATATAT 4481

RESULT 3  
US-09-707-919A-20  
; Sequence 20, Application US/09707919A  
; Patent No. 6623927  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: Method for detection of human spinocerebellar ataxia 2  
; FILE OF INVENTION: gene variants  
; FILE REFERENCE: US 443  
; CURRENT APPLICATION NUMBER: US/09/707,919A  
; CURRENT FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 4481  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-707-919A-20

Query Match 100.0%; Score 4481; DB 3; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCTCTCTACAGTGTCTCCCGCGGCCCGGG 60  
Db 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCTCTCTACAGTGTCTCCCGCGGCCCGGG 60

Qy 61 GCCACCTCAAGTTCTGCTTCCGTCTGACCCCTCCGACTTCCGTTAAAGAGTCCCTATCCG 120  
Db 61 GCCACCTCAAGTTCTGCTTCCGTCTGACCCCTCCGACTTCCGTTAAAGAGTCCCTATCCG 120  
Qy 121 CACCTTCGCTCCCAACCCGCGGCTCCGCGCGCCCTCCGATGCGCTCAGCGCCGCCA 180  
Db 121 CACCTTCGCTCCCAACCCGCGGCTCCGCGCGCCCTCCGATGCGCTCAGCGCCGCCA 180  
Qy 181 GCTCTCGAGTCCCGGCTGGCCACCGAGTCTTCGCGCTTCGCGCACAGGAGTGGGCC 240  
Db 181 GCTCTCGAGTCCCGGCTGGCCACCGAGTCTTCGCGCTTCGCGCACAGGAGTGGGCC 240  
Qy 241 GGTGCGCTCGCTCCAGCGGCTGGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGG 300  
Db 241 GGTGCGCTCGCTCCAGCGGCTGGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGG 300  
Qy 301 CCGGACCGTATCCCTCCGCGGCTCCGCGCGCCCTCCGCGCGGCGGCGGCGGCGG 360  
Db 301 CCGGACCGTATCCCTCCGCGGCTCCGCGCGCCCTCCGCGCGGCGGCGGCGGCGG 360  
Qy 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCGGCGGCGGCGGCG 420  
Db 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCGGCGGCGGCGGCG 420  
Qy 421 TTTGCGCGCGCTCCGCGGCTCTTGGTCTCGGCGGCGCTCCCGCGCCCTTCTGTCGTC 480  
Db 421 TTTGCGCGCGCTCCGCGGCTCTTGGTCTCGGCGGCGCTCCCGCGCCCTTCTGTCGTC 480  
Qy 481 GTCTTCTCCCTCCGCGGCTCCGCGGCTCCGCGCGGCGGCGGCGGCGGCGGCGG 540  
Db 481 GTCTTCTCCCTCCGCGGCTCCGCGGCTCCGCGCGGCGGCGGCGGCGGCGGCGG 540  
Qy 541 CTCGCGCGCGCTGCTGCTCCGCGGCTCCGCGGCTCTCTTGGCGGCGGCGGCGGCG 600  
Db 541 CTCGCGCGCGCTGCTGCTCCGCGGCTCCGCGGCTCTCTTGGCGGCGGCGGCGGCG 600  
Qy 601 TGTCCCGCGCGGCTGCGAGCGGCTGTATGGGCGCTTCCACATGTCGCTGAAGCGGCG 660  
Db 601 TGTCCCGCGCGGCTGCGAGCGGCTGTATGGGCGCTTCCACATGTCGCTGAAGCGGCG 660  
Qy 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720  
Db 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720  
Qy 721 CAGCGCGCGCGGCGGCTGCGCAATGTCCGCAAGCGCGGCGGCGGCGGCGGCGGCT 780  
Db 721 CAGCGCGCGCGGCGGCTGCGCAATGTCCGCAAGCGCGGCGGCGGCGGCGGCT 780  
Qy 781 CCGCGCGCGCGGCTGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 CCGCGCGCGCGGCTGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Qy 841 TCCTCGGTGCTGCGGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900  
Db 841 TCCTCGGTGCTGCGGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900  
Qy 901 AGTAACAAAGGAGTCCCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAAAATAGG 960  
Db 901 AGTAACAAAGGAGTCCCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAAAATAGG 960  
Qy 961 ATGTTTCATATACCTTACATGATTTGCTGCTCCAAATGTGAAGTACAAGTAAAAATGA 1020  
Db 961 ATGTTTCATATACCTTACATGATTTGCTGCTCCAAATGTGAAGTACAAGTAAAAATGA 1020  
Qy 1021 GGTATATATGAAGGAGTAAATTTTAAACCTTACAGTCCGAGTGTGATTTGCTGCTGCT 1080  
Db 1021 GGTATATATGAAGGAGTAAATTTTAAACCTTACAGTCCGAGTGTGATTTGCTGCTGCT 1080  
Qy 1081 GCACATGAGAAAGTACAGAAATCCAGTTTCGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
Db 1081 GCACATGAGAAAGTACAGAAATCCAGTTTCGCGGCGGCGGCGGCGGCGGCGGCGG 1140  
Qy 1141 ATTTGTTCAAAATGTTTTCAGACTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200

|||||  
1141 ATTTTGTCAAATGTTACAGACTTTGTTGGTACAGTTTAAAGATATGACTCCAGTTAT 1200  
QY  
1201 GCAAAAAGAGATGCTTTTAACTGACTCTGCTATCAGTGTAAAGTGAATGGCGAAACACAAA 1260  
Db  
1201 GCAAAAAGAGATGCTTTTAACTGACTCTGCTATCAGTGTAAAGTGAATGGCGAAACACAAA 1260  
QY  
1261 GAGAAGGACCTGGAGCCCTGGGATGAGGTGAATCTCAGACCAATGAGGAACTTGAGGCT 1320  
Db  
1261 GAGAAGGACCTGGAGCCCTGGGATGAGGTGAATCTCAGACCAATGAGGAACTTGAGGCT 1320  
QY  
1321 TTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380  
Db  
1321 TTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380  
QY  
1381 AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440  
Db  
1381 AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440  
QY  
1441 AGAGATAAATCTCAGAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA 1500  
Db  
1441 AGAGATAAATCTCAGAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA 1500  
QY  
1501 ATTTAGTCAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAAAATGATATAGGAGTGAG 1560  
Db  
1501 ATTTAGTCAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAAAATGATATAGGAGTGAG 1560  
QY  
1561 GAAAGAAAATACACAGCAGTTTACAGAGAAATTTCCAGTGAACGTGAGGGGACACAGCATAAAC 1620  
Db  
1561 GAAAGAAAATACACAGCAGTTTACAGAGAAATTTCCAGTGAACGTGAGGGGACACAGCATAAAC 1620  
QY  
1621 ACTAGGGAAAATAAATATATATCTCTCGGACAAAGAAAATAGAGAGTATATCTCTGGGA 1680  
Db  
1621 ACTAGGGAAAATAAATATATATCTCTCGGACAAAGAAAATAGAGAGTATATCTCTGGGA 1680  
QY  
1681 AGTGGAGACAGAAATTCACCGGTATGGGCGAGCCTGGATCGGGTCCATGCCATCAAGA 1740  
Db  
1681 AGTGGAGACAGAAATTCACCGGTATGGGCGAGCCTGGATCGGGTCCATGCCATCAAGA 1740  
QY  
1741 TCCACTTCTCACACTTTCAGATTTCAACCCGAAATCTGGTTTCAGACCAAGAGTAGTTAAT 1800  
Db  
1741 TCCACTTCTCACACTTTCAGATTTCAACCCGAAATCTGGTTTCAGACCAAGAGTAGTTAAT 1800  
QY  
1801 GGAAGTGTTCCTCGGCCATCGCCTTGCCCATCTCTTCCTCTCGGCCACCTTCTCGCTAC 1860  
Db  
1801 GGAAGTGTTCCTCGGCCATCGCCTTGCCCATCTCTTCCTCTCGGCCACCTTCTCGCTAC 1860  
QY  
1861 CAGTCAGTCCCAACTCTCTTCCACTCGGGAGCCACCCCTACACGGCGGCCCTCAGG 1920  
Db  
1861 CAGTCAGTCCCAACTCTCTTCCACTCGGGAGCCACCCCTACACGGCGGCCCTCAGG 1920  
QY  
1921 CCCCCCTCGGGGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGCTCCT 1980  
Db  
1921 CCCCCCTCGGGGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGCTCCT 1980  
QY  
1981 GTCTCTATATGCTTAACGCATGTCTTTCAGAGGGCCCTCCAAAGGATGTCCCAAAGGCC 2040  
Db  
1981 GTCTCTATATGCTTAACGCATGTCTTTCAGAGGGCCCTCCAAAGGATGTCCCAAAGGCC 2040  
QY  
2041 CAGCGACATCCTCGAAATACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100  
Db  
2041 CAGCGACATCCTCGAAATACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100  
QY  
2101 GAAATTTGATCCCAACACCCAGTGAAGCAGTACTCTCCAGTGTAGCAAGGACAGT 2160  
Db  
2101 GAAATTTGATCCCAACACCCAGTGAAGCAGTACTCTCCAGTGTAGCAAGGACAGT 2160  
QY  
2161 CCCTCGGGGGAAAGTGGTTCATCAGTGGTTCAGTGGGTTTCCAAAGATTTATCCCCATAAAT 2220  
Db  
2161 CCCTCGGGGGAAAGTGGTTCATCAGTGGTTCAGTGGGTTTCCAAAGATTTATCCCCATAAAT 2220  
QY  
2221 CATAGACCCAGGTTCTCCGACAGACAGCAGTATTTGGAAAATACCCCGAGTGGGCCAGTTCTT 2280  
|||||

Db  
2221 CATAGACCCAGGTTCTCCGACAGACAGATATTGAAAATACCCCGAGTGGGCCAGTTCTT 2280  
QY  
2281 GCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTTGCCATGCTTATTCAGCTGCA 2340  
Db  
2281 GCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTTGCCATGCTTATTCAGCTGCA 2340  
QY  
2341 TCTCCTAGCCTGCTAGTCTGCTGCAACAGAGCTGTACCCTTCTAGTCAGGCTAAA 2400  
Db  
2341 TCTCCTAGCCTGCTAGTCTGCTGCAACAGAGCTGTACCCTTCTAGTCAGGCTAAA 2400  
QY  
2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGAGGGAATAAAGAAAATATATAA 2460  
Db  
2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGAGGGAATAAAGAAAATATATAA 2460  
QY  
2461 CCCAATGAAAATCATCCTAGCTTCTCAAAGCTGAAAACAAAGGTATATCAGCAGTTGT 2520  
Db  
2461 CCCAATGAAAATCATCCTAGCTTCTCAAAGCTGAAAACAAAGGTATATCAGCAGTTGT 2520  
QY  
2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTTAGTTTACAG 2580  
Db  
2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTTAGTTTACAG 2580  
QY  
2581 CCAAGTCTTACTTCTGAAATCTATGGATCAACTTACTATAAACAAGGATAGAGGAGAGAAA 2640  
Db  
2581 CCAAGTCTTACTTCTGAAATCTATGGATCAACTTACTATAAACAAGGATAGAGGAGAGAAA 2640  
QY  
2641 TCAAGAGATTTGATCAAAGACAAAATTTGAACCAAGTGTCTAAGGATTTCTTCAATGAAAAT 2700  
Db  
2641 TCAAGAGATTTGATCAAAGACAAAATTTGAACCAAGTGTCTAAGGATTTCTTCAATGAAAAT 2700  
QY  
2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCGAATAGCCCGAGCATTTCCCTTCA 2760  
Db  
2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCGAATAGCCCGAGCATTTCCCTTCA 2760  
QY  
2761 ATACTTAGTAAACACGGAGCACAAAGGGGACCTGAGTCACTTCCCAAGGGGTTTCAAGCT 2820  
Db  
2761 ATACTTAGTAAACACGGAGCACAAAGGGGACCTGAGTCACTTCCCAAGGGGTTTCAAGCT 2820  
QY  
2821 TCCAGCCAGCATGTAAAACAGAGAAAGACGATAAGGAAGAGAAAGACGAGCTGAG 2880  
Db  
2821 TCCAGCCAGCATGTAAAACAGAGAAAGACGATAAGGAAGAGAAAGACGAGCTGAG 2880  
QY  
2881 CAAAGTGAAGAAATCAACATTGAAATCCCAATGAAAGGAGTTCAACCCAGTTCCTCTCT 2940  
Db  
2881 CAAAGTGAAGAAATCAACATTGAAATCCCAATGAAAGGAGTTCAACCCAGTTCCTCTCT 2940  
QY  
2941 CAGCCAAAGCCTTCTACTACCCCACTTCCACCTCGSCCTCAAGACAACTAGCCCATCT 3000  
Db  
2941 CAGCCAAAGCCTTCTACTACCCCACTTCCACCTCGSCCTCAAGACAACTAGCCCATCT 3000  
QY  
3001 ATGGTGGGTCAATCAACAGCCTTCTCAGTCTTATCTCAGCCTGTTTGTGTCACCAAT 3060  
Db  
3001 ATGGTGGGTCAATCAACAGCCTTCTCAGTCTTATCTCAGCCTGTTTGTGTCACCAAT 3060  
QY  
3061 ATGATGTATCCAGTCCAGTGGCCAGCGGTGCAACCTTTATACCAATACCTATGACG 3120  
Db  
3061 ATGATGTATCCAGTCCAGTGGCCAGCGGTGCAACCTTTATACCAATACCTATGACG 3120  
QY  
3121 CCCATGCCAGTGAATCAAGCAGACATATAGAGCAGTACCATAATGCCCCAACAGCGG 3180  
Db  
3121 CCCATGCCAGTGAATCAAGCAGACATATAGAGCAGTACCATAATGCCCCAACAGCGG 3180  
QY  
3181 CAAGACACGATCATCAGAGTGCATGATGACCCAGCGTTCAGCAGCGGCCCGCCAGTT 3240  
Db  
3181 CAAGACACGATCATCAGAGTGCATGATGACCCAGCGTTCAGCAGCGGCCCGCCAGTT 3240  
QY  
3241 GAGCCACCCCGCCAGCTTACTCCAGCAATATGTTGGCTACAGTCTCAGAGTTCCCA 3300  
Db  
3241 GAGCCACCCCGCCAGCTTACTCCAGCAATATGTTGGCTACAGTCTCAGAGTTCCCA 3300  
QY  
3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTACGTCTCAGCATCTCATGTCTATAGT 3360  
Db  
3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTACGTCTCAGCATCTCATGTCTATAGT 3360

3361 CCTGTAATACAGGGTAATGCTAGATGATGGACACCAACACAGCCCGCCTGGTTTA 3420  
3361 CCTGTAATACAGGGTAATGCTAGATGATGGACACCAACACAGCCCGCCTGGTTTA 3420  
3421 GTATCTTTCTTCAGCAACTCAGTACGGGGCTCATGACGACGCGATGCGATGATCATGT 3480  
3421 GTATCTTTCTTCAGCAACTCAGTACGGGGCTCATGACGACGCGATGCGATGATCATGT 3480  
3481 CCCAAATTACCATAACAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540  
3481 CCCAAATTACCATAACAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540  
3541 TCCCTTGTCTCAGCAGTATGCGCACCCCTAAACGCTACCTTGACCCACATATCTCCACACCT 3600  
3541 TCCCTTGTCTCAGCAGTATGCGCACCCCTAAACGCTACCTTGACCCACATATCTCCACACCT 3600  
3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCCAACATGGTGGAGTCACTCTGCA 3660  
3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCCAACATGGTGGAGTCACTCTGCA 3660  
3661 CCAGTCTCTGTTACGACCATCAGCACAGGCGCCGAGGCTCTCCATCTGSCCAGTCCA 3720  
3661 CCAGTCTCTGTTACGACCATCAGCACAGGCGCCGAGGCTCTCCATCTGSCCAGTCCA 3720  
3721 CAGCAGCAGTACGCACTTTACACCGGGGCTTGGCCCAACTCCACCTCCATGACACCT 3780  
3721 CAGCAGCAGTACGCACTTTACACCGGGGCTTGGCCCAACTCCACCTCCATGACACCT 3780  
3781 GCCTCCACAGCAGTCCGACAGATAGTTTCCAGCAGCAACACAGACTGCTTTAGC 3840  
3781 GCCTCCACAGCAGTCCGACAGATAGTTTCCAGCAGCAACACAGACTGCTTTAGC 3840  
3841 ATCCATCTCTTCAGTTCAGCGCGGTATACCAACCCACATGCGCCACGCTACT 3900  
3841 ATCCATCTCTTCAGTTCAGCGCGGTATACCAACCCACATGCGCCACGCTACT 3900  
3901 CAGGCTCATGTACAGTCAAGATGTTCTTCTCATCAACTGCCCATGCCCAATGATG 3960  
3901 CAGGCTCATGTACAGTCAAGATGTTCTTCTCATCAACTGCCCATGCCCAATGATG 3960  
3961 CTAATGACGACACGCGCCGCGTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020  
3961 CTAATGACGACACGCGCCGCGTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020  
4021 CCATTCAGTCTCGACAAACAGCGCATTTCCCTATATGACGACCCCTTCAGTCAAGCC 4080  
4021 CCATTCAGTCTCGACAAACAGCGCATTTCCCTATATGACGACCCCTTCAGTCAAGCC 4080  
4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGGAGGAAACCGAAAGGCCAAATTCCTCCTC 4140  
4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGGAGGAAACCGAAAGGCCAAATTCCTCCTC 4140  
4141 CTTTCTACTGCTTACCAACTGGAAGCACAGAAACTAGAAATTCATTTATTTGTTTT 4200  
4141 CTTTCTACTGCTTACCAACTGGAAGCACAGAAACTAGAAATTCATTTATTTGTTTT 4200  
4201 TAAATATATATGTTGATTTCTTTGTAAACATCAATAGGAATGCTAACAGTTTCACTTGAG 4260  
4201 TAAATATATATGTTGATTTCTTTGTAAACATCAATAGGAATGCTAACAGTTTCACTTGAG 4260  
4261 TGGAAAGATCTTGGACCGAGTAGAGGCATTTTAGGAACTTGGGGCTATTTCATATTTCCA 4320  
4261 TGGAAAGATCTTGGACCGAGTAGAGGCATTTTAGGAACTTGGGGCTATTTCATATTTCCA 4320  
4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380  
4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380  
4381 ATTTTTTAAACCCCTTGAAGTCTATGAACATCAGCTTAGCAAAAGAAAGTAAACAGAGT 4440  
4381 ATTTTTTAAACCCCTTGAAGTCTATGAACATCAGCTTAGCAAAAGAAAGTAAACAGAGT 4440

QY 4441 GATTCTTGCTCTACTTACTGCTATAAAAAAAAAAAAAAAAAAAAA 4481  
Db |||||  
QY 4441 GATTCTTGCTCTACTTACTGCTATAAAAAAAAAAAAAAAAAAAAA 4481  
Db |||||  
RESULT 4  
US-09-083-268-2  
; Sequence 2, Application US/09083268  
; Patent No. 6673535  
; GENERAL INFORMATION:  
; APPLICANT: Pulst, Stefan M  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR  
; TITLE OF INVENTION: ATAXIA-2 AND PRODUCTS RELATED THERETO  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Muetling, Raasch & Gebhardt, P.A.  
; STREET: 119 No. 6673535th Fourth Street  
; CITY: Minneapolis  
; STATE: Minnesota  
; COUNTRY: USA  
; ZIP: 55401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/083,268  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/727,084  
; FILING DATE: 08-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McCormack, Myra H  
; REGISTRATION NUMBER: 36,602  
; REFERENCE/DOCKET NUMBER: 232.00010101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/305-1220  
; TELEFAX: 612/305-1228  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4481 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 163..4101  
US-09-083-268-2

Query Match 100.0%; Score 4481; DB 3; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCCCGCTCCCTCACGTGCTCCCTCCCGGCCCCGGG 60  
Db 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCCCGCTCCCTCACGTGCTCCCTCCCGGCCCCGGG 60  
QY 61 GCCACCTCACTGCTTGTGTTCCGTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
Db 61 GCCACCTCACTGCTTGTGTTCCGTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
QY 121 CACCTCCGCTCCACCCCGCGCTCCGGCGCGCCCTCCGATCGCGTCAAGCGGCCCA 180  
Db 121 CACCTCCGCTCCACCCCGCGCTCCGGCGCGCCCTCCGATCGCGTCAAGCGGCCCA 180  
QY 181 GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTCGCGCAGCAGGTTGGCCC 240  
Db 181 GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTCGCGCAGCAGGTTGGCCC 240  
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGCGGGCGCGGTGGCGCGGCC 300



QY 2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCAACAGTTGTT 2520  
Db 2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCAACAGTTGTT 2520  
QY 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTAGTTACAG 2580  
Db 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTAGTTACAG 2580  
QY 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAACAAAAATAGAGAGGAGAAAAA 2640  
Db 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAACAAAAATAGAGAGGAGAAAAA 2640  
QY 2641 TCAAGAGATTTGATCAAGAGAAATTTGAAACCAAGTGTAAAGATTTCTTTCAATGAAAT 2700  
Db 2641 TCAAGAGATTTGATCAAGAGAAATTTGAAACCAAGTGTAAAGATTTCTTTCAATGAAAT 2700  
QY 2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAGCCGAATAGCCAGCATTTCCCTTCA 2760  
Db 2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAGCCGAATAGCCAGCATTTCCCTTCA 2760  
QY 2761 ATACTTAGTAAACACGGAGCAAGAGGGGACCTGAGGTCACTTCCAAAGGGGTTCAAGACT 2820  
Db 2761 ATACTTAGTAAACACGGAGCAAGAGGGGACCTGAGGTCACTTCCAAAGGGGTTCAAGACT 2820  
QY 2821 TCCAGCCGACGATGTAACAGAGAAAGACGATAGGAAGAGAAAGACGAGCTGAG 2880  
Db 2821 TCCAGCCGACGATGTAACAGAGAAAGACGATAGGAAGAGAAAGACGAGCTGAG 2880  
QY 2881 CAAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCT 2940  
Db 2881 CAAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCT 2940  
QY 2941 CAGCCAAAGCCCTTCTACTACCCCAACTTCACTCGGCTCAAGCAACACTAGCCCATCT 3000  
Db 2941 CAGCCAAAGCCCTTCTACTACCCCAACTTCACTCGGCTCAAGCAACACTAGCCCATCT 3000  
QY 3001 ATGGTGGGTATCAACAGCAAACTCCAGTTTATCTCAGCCCTGTTGTTTGCACCAAT 3060  
Db 3001 ATGGTGGGTATCAACAGCAAACTCCAGTTTATCTCAGCCCTGTTGTTTGCACCAAT 3060  
QY 3061 ATGATGTATCCAGTCCAGTGGCCGAGCGGTGCAACCTTTATACCCAATACCTATGAG 3120  
Db 3061 ATGATGTATCCAGTCCAGTGGCCGAGCGGTGCAACCTTTATACCCAATACCTATGAG 3120  
QY 3121 CCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAAAATATGCCCAACAGCGG 3180  
Db 3121 CCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAAAATATGCCCAACAGCGG 3180  
QY 3181 CAAAGCCAGCATCATCAGATGCCATGATGCAACCCAGCGTCAGCAGCGGGCCACCGATT 3240  
Db 3181 CAAAGCCAGCATCATCAGATGCCATGATGCAACCCAGCGTCAGCAGCGGGCCACCGATT 3240  
QY 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTTACAGTCTCAGCAGTTCCCA 3300  
Db 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTTACAGTCTCAGCAGTTCCCA 3300  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCAATATCAGTCTCAGCATCTCATGTCTATAGT 3360  
Db 3301 AATCAGCCCTTGTTCAGCATGTGCCAATATCAGTCTCAGCATCTCATGTCTATAGT 3360  
QY 3361 CTTGTAATACAGGGTAATGCTAGATGATGGCACCACCAACACACACGCGGCTGGTTTA 3420  
Db 3361 CTTGTAATACAGGGTAATGCTAGATGATGGCACCACCAACACACACGCGGCTGGTTTA 3420  
QY 3421 GTATCTTTCTCAGCAACTCAGTACGGGCTCATGAGCAGCGCATGCGATGATGCTATG 3480  
Db 3421 GTATCTTTCTCAGCAACTCAGTACGGGCTCATGAGCAGCGCATGCGATGATGCTATG 3480  
QY 3481 CCCAAATACCATCAACAAAGGAGCAAGCCCTTCTTCTACTTTGCCATTTCCACGGG 3540  
Db 3481 CCCAAATACCATCAACAAAGGAGCAAGCCCTTCTTCTACTTTGCCATTTCCACGGG 3540

QY 3541 TCCCTTGCTCAGCAGTATGCCGCAACCTTAAACGCTACCTTGCAACCACATATCTTCCACACCT 3600  
Db 3541 TCCCTTGCTCAGCAGTATGCCGCAACCTTAAACGCTACCTTGCAACCACATATCTTCCACACCT 3600  
QY 3601 CAGCCTTCAGCTACCCACCTGGACAGCAGCAAGCCAAATGTTGGAAGTCATCTTGCAC 3660  
Db 3601 CAGCCTTCAGCTACCCACCTGGACAGCAGCAAGCCAAATGTTGGAAGTCATCTTGCAC 3660  
QY 3661 CCCAGTCTCTGTTACGACCATCAGCAGCAGCCGCGCCAGGCTCTCCATCTGGCCAGTCCA 3720  
Db 3661 CCCAGTCTCTGTTACGACCATCAGCAGCAGCCGCGCCAGGCTCTCCATCTGGCCAGTCCA 3720  
QY 3721 CAGCAGCAGTCAAGCATTTTACACGCGGGGCTTGGCCAACTCCACCCCTCCATGACACT 3780  
Db 3721 CAGCAGCAGTCAAGCATTTTACACGCGGGGCTTGGCCAACTCCACCCCTCCATGACACT 3780  
QY 3781 GCCTCCAAACGCGAGTCGSCACAGAAATAGTTTCCACGAGCAGCAACAGACTGCTTTACG 3840  
Db 3781 GCCTCCAAACGCGAGTCGSCACAGAAATAGTTTCCACGAGCAGCAACAGACTGCTTTACG 3840  
QY 3841 ATCCATCCTTCTCAGCTTTCAGCCGCGGTATACCAACCCACATGCGGCCACGTACT 3900  
Db 3841 ATCCATCCTTCTCAGCTTTCAGCCGCGGTATACCAACCCACATGCGGCCACGTACT 3900  
QY 3901 CAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACTGCCCAATGATG 3960  
Db 3901 CAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACTGCCCAATGATG 3960  
QY 3961 CTAATGAGCAGCAGCCACCGCGGTCCCGAGCGCCCTCGCTCAAGTGCATACAG 4020  
Db 3961 CTAATGAGCAGCAGCCACCGCGGTCCCGAGCGCCCTCGCTCAAGTGCATACAG 4020  
QY 4021 CCCATTCCAGTCTCGAACAACAGCGCATTTCCCTTATATGACGCAACCTTTCAGTACAAGCC 4080  
Db 4021 CCCATTCCAGTCTCGAACAACAGCGCATTTCCCTTATATGACGCAACCTTTCAGTACAAGCC 4080  
QY 4081 CACCACCAACAGCAGTTGTAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140  
Db 4081 CACCACCAACAGCAGTTGTAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140  
QY 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGCAAGAACTAGAAATTTTATTTTGTGTTT 4200  
Db 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGCAAGAACTAGAAATTTTATTTTGTGTTT 4200  
QY 4201 TAAATATATATGTTGATTTCTTGTAAACATCCAATAGGAATGCTAAACAGTTTCACTTCGAG 4260  
Db 4201 TAAATATATATGTTGATTTCTTGTAACTCAATAGGAATGCTAAACAGTTTCACTTCGAG 4260  
QY 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAATTTGGGGCTATTCCATTAATTTCCA 4320  
Db 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAATTTGGGGCTATTCCATAATTTCCA 4320  
QY 4321 TATGCTGTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTTATTT 4380  
Db 4321 TATGCTGTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTTATTT 4380  
QY 4381 ATTTTTTAATACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAAGAAAGTAAACAAGT 4440  
Db 4381 ATTTTTTAATACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAAGAAAGTAAACAAGT 4440  
QY 4441 GATTCTTGCTGCTATTTACTGCTAAAAAATAAAAAA 4481  
Db 4441 GATTCTTGCTGCTATTTACTGCTAAAAAATAAAAAA 4481

## RESULT 5

US-08-981-998A-2

; Sequence 2, Application US/08981998A

; Patent No. 684431

; GENERAL INFORMATION:

; APPLICANT: PULST, STEPHAN M.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR

; ATAXIA-2 AND PRODUCTS RELATED THERETO



NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSER: MUEITING, RAASCH & GEBHARDT, P.A.  
STREET: 119 NORTH FOURTH STREET, SUITE 203  
CITY: MINNEAPOLIS  
STATE: MINNESOTA  
COUNTRY: 55401  
ZIP: 55401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/981,998A  
FILING DATE: 11-May-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 97/42314  
FILING DATE: 08-MAY-1997  
APPLICATION NUMBER: US 08/727,084  
FILING DATE: 08-OCT-1996  
APPLICATION NUMBER: US 60/022,207  
FILING DATE: 19-JUL-1996  
APPLICATION NUMBER: US 60/017,388  
FILING DATE: 08-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: MUEITING, ANN M.  
REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 232.00010120  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-305-1217  
TELEFAX: 612-305-1228  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4481 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 163...4101  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-08-981-998A-2

Query Match 100.0%; Score 4481; DB 3; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCTCCTCAGCTGTCCCTCCCGGCCCGGG 60  
DB 1 ACCCCGAGAAAGCAACCCAGCGCGCGCTCCTCAGCTGTCCCTCCCGGCCCGGG 60  
QY 61 GCACCTCAGCTTCTGCTCCGCTGTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
DB 61 GCACCTCAGCTTCTGCTCCGCTGTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
QY 121 CACCTCGGCTCCACCCGGCGCTCCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA 180  
DB 121 CACCTCGGCTCCACCCGGCGCTCCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA 180  
QY 181 GCTCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCGAGCGAGTGGCCC 240  
DB 181 GCTCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCGAGCGAGTGGCCC 240  
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
DB 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
QY 301 CCGGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
DB 301 CCGGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360

QY 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGTGTAGCAACGGCAACGGCGCGCGCG 420  
DB 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGTGTAGCAACGGCAACGGCGCGCGCG 420  
QY 421 TTTCCGCGCGCGCTCCCGCGCGCTCCTTGTGTCTCGCGCGCGCTCCCGCGCGCTTCTGCTC 480  
DB 421 TTTCCGCGCGCGCTCCCGCGCGCTCCTTGTGTCTCGCGCGCGCTCCCGCGCGCTTCTGCTC 480  
QY 481 GTCTCTTCTCCCTCCGCTCAGCGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCTCCCGCG 540  
DB 481 GTCTCTTCTCCCTCCGCTCAGCGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCTCCCGCG 540  
QY 541 CTGCGCGCGCTGCTGCTCCCGCGCGCTTCCGCGCTTCTGCGCGCGCTCCCGCGCGCTCCCGCG 600  
DB 541 CTGCGCGCGCTGCTGCTCCCGCGCGCTTCCGCGCTTCTGCGCGCGCTCCCGCGCGCTCCCGCG 600  
QY 601 TGTCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 601 TGTCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 661 CAG 720  
DB 661 CAG 720  
QY 721 CAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 721 CAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
QY 781 CCGCGCGCGCGCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 781 CCGCGCGCGCGCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 TCCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 841 TCCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 901 AGTAAACAAAGGACTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 901 AGTAAACAAAGGACTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
QY 961 ATGTTTCATATATACATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 961 ATGTTTCATATATACATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 GGTATATATGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
DB 1021 GGTATATATGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCG 1140  
DB 1081 GCACATGAGAAAGTACAGAAATCCAGTTCG 1140  
QY 1141 ATTTTGTTCAAATGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
DB 1141 ATTTTGTTCAAATGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1201 GCAAAAGAGATGCTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
DB 1201 GCAAAAGAGATGCTTTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
QY 1261 GAGAAAGCCTCGAGCGCTGGGATGAGTGAATCTCAGACCGCAATGAGGAACTTTGAGGCT 1320  
DB 1261 GAGAAAGCCTCGAGCGCTGGGATGAGTGAATCTCAGACCGCAATGAGGAACTTTGAGGCT 1320  
QY 1321 TTGGAATAATGACGTATCTAATGATGGATCCCAATGATATGTTTCGATATATATGAGAA 1380  
DB 1321 TTGGAATAATGACGTATCTAATGATGGATCCCAATGATATGTTTCGATATATATGAGAA 1380  
QY 1381 AATTATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTTAGAA 1440  
DB 1381 AATTATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTTAGAA 1440

QY	1441	AGAGATAACTCAGAGAATTTTTTAAACCGGAAGCAAGGGCAAAACAGGTTAGCAGAGAA	1500	Db	2521	TCTGAACATAGAAAAACAGATTGATGATTTAAAGAAATTTAAAGATGATTTTATAGGTTACAG	2580
Db	1441	AGAGATAACTCAGAGAATTTTTTAAACCGGAAGCAAGGGCAAAACAGGTTAGCAGAGAA	1500	QY	2581	CAAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAATAGAGAGGAGAAAA	2640
QY	1501	ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATGATAGGAGTGAG	1560	Db	2581	CAAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAATAGAGAGGAGAAAA	2640
Db	1501	ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATGATAGGAGTGAG	1560	QY	2641	TCAAGAGATTTGATCAAGACAAAAATTTGAACAAAGTGTCTAAGGATTTCTTTCAATTGAAAAAT	2700
QY	1561	GAAGAAAAATACACAGCAGTTCCAGAGAAATTCAGTGAACCTGAGGGGCACAGCATAAAC	1620	Db	2641	TCAAGAGATTTGATCAAGACAAAAATTTGAACAAAGTGTCTAAGGATTTCTTTCAATTGAAAAAT	2700
Db	1561	GAAGAAAAATACACAGCAGTTCCAGAGAAATTCAGTGAACCTGAGGGGCACAGCATAAAC	1620	QY	2701	AGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCGGAATAGCCCCAGCATTTTCCCTTTCA	2760
QY	1621	ACTAGGGAATAAATATATATCTCTCTGGAACAAAGAAATAGAGAAAGTCAATATCTCTGGGA	1680	Db	2701	AGCAGCAGCAACTGTGTACAGTGGCAGCAGCAAGCGGAATAGCCCCAGCATTTTCCCTTTCA	2760
Db	1621	ACTAGGGAATAAATATATATCTCTCTGGAACAAAGAAATAGAGAAAGTCAATATCTCTGGGA	1680	QY	2761	ATACTTTAGTAACCGGAGCAAGAGGGGACCTCAGGTTCACCTCCCAAGGGGTTTCAGACT	2820
QY	1681	AGTGGGAGACAAATTCACCGGTATGGCCAGCCTGGATCGGGTCCATGSCATCAAGA	1740	Db	2761	ATACTTTAGTAACCGGAGCAAGAGGGGACCTCAGGTTCACCTCCCAAGGGGTTTCAGACT	2820
Db	1681	AGTGGGAGACAAATTCACCGGTATGGCCAGCCTGGATCGGGTCCATGSCATCAAGA	1740	QY	2821	TCCAGCCCGCATGTAAACAAAGAGAACGATTAAGGAAGAGAAAGACGCGAGCTGAG	2880
QY	1741	TCCACTTCTCACATTCAGATTTCAACCCCGAATTCCTGGTTCAGACCAAGAGTAGTTAAT	1800	Db	2821	TCCAGCCCGCATGTAAACAAAGAGAACGATTAAGGAAGAGAAAGACGCGAGCTGAG	2880
Db	1741	TCCACTTCTCACATTCAGATTTCAACCCCGAATTCCTGGTTCAGACCAAGAGTAGTTAAT	1800	QY	2881	CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGGGAGTTCAACCCACGTTTCTCTCT	2940
QY	1801	GGAGTGTTCCTGGCCATCGCCTTGGCCATCTCCTCTCGCCACCTTCTCGCTAC	1860	Db	2881	CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGGGAGTTCAACCCACGTTTCTCTCT	2940
Db	1801	GGAGTGTTCCTGGCCATCGCCTTGGCCATCTCCTCTCGCCACCTTCTCGCTAC	1860	QY	2941	CAGCAAGCCCTTCTACTACCCCAACTTTCACCTCGGCTCAAGCACAACCTAGGCCATCT	3000
QY	1861	CAGTCAGTCCCACTCTCTCCACTCGGGCAGCCACCCTACACGGCCGCCCTCCAGG	1920	Db	2941	CAGCAAGCCCTTCTACTACCCCAACTTTCACCTCGGCTCAAGCACAACCTAGGCCATCT	3000
Db	1861	CAGTCAGTCCCACTCTCTCCACTCGGGCAGCCACCCTACACGGCCGCCCTCCAGG	1920	QY	3001	ATGGTGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGCCCTGTTGTTTGCACCAAT	3060
QY	1921	CCCCCTCGCGCCCATCAGACCCCCGCTCTACCCCTCTGCTCATGTTCTCCAGCTCCT	1980	Db	3001	ATGGTGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGCCCTGTTGTTTGCACCAAT	3060
Db	1921	CCCCCTCGCGCCCATCAGACCCCCGCTCTACCCCTCTGCTCATGTTCTCCAGCTCCT	1980	QY	3061	ATGATGTATCCAGTCCAGTGAGCCAGCGTGCAACTTTATACCCCAATACCTATGAGC	3120
QY	1981	GTCTCTACTATGCCTTAAAGCAGTGTCTTCAGAGGGCCTCCAGGATGTCCCCAAAGGCC	2040	Db	3061	ATGATGTATCCAGTCCAGTGAGCCAGCGTGCAACTTTATACCCCAATACCTATGAGC	3120
Db	1981	GTCTCTACTATGCCTTAAAGCAGTGTCTTCAGAGGGCCTCCAGGATGTCCCCAAAGGCC	2040	QY	3121	CCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCATAATATGCCCCAACAGCGG	3180
QY	2041	CAGCGACATCTTCGAAATCACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCCTA	2100	Db	3121	CCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCATAATATGCCCCAACAGCGG	3180
Db	2041	CAGCGACATCTTCGAAATCACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCCTA	2100	QY	3181	CAAGCAGCAGCATCATCAGAGTGCCATGATGCAACCCAGCGTCAGCAGCGGGCCCCAGGAT	3240
QY	2101	GAAATTTGATCCCAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGGACCAGT	2160	Db	3181	CAAGCAGCAGCATCATCAGAGTGCCATGATGCAACCCAGCGTCAGCAGCGGGCCCCAGGAT	3240
Db	2101	GAAATTTGATCCCAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGGACCAGT	2160	QY	3241	GCAGCCACCCACAGCTTACTCCACGCAATATGTTGCTTACAGTCTCAGCAGTTCCTCA	3300
QY	2161	CCCTCGGGGGNAACGTGTGCATCAGTGTGAGTGGGGTTCCAGATTTATCCCTTAAACT	2220	Db	3241	GCAGCCACCCACAGCTTACTCCACGCAATATGTTGCTTACAGTCTCAGCAGTTCCTCA	3300
Db	2161	CCCTCGGGGGNAACGTGTGCATCAGTGTGAGTGGGGTTCCAGATTTATCCCTTAAACT	2220	QY	3301	AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCCTCATGTCTATAGT	3360
QY	2221	CATAGACCAGGTCTCCAGACAGAAACAGTATGGAATATCCCGAGTGGCCAGTTCCT	2280	Db	3301	AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCCTCATGTCTATAGT	3360
Db	2221	CATAGACCAGGTCTCCAGACAGAAACAGTATGGAATATCCCGAGTGGCCAGTTCCT	2280	QY	3361	CCTGTAATACAGGGTAATGCTTAGAATGATGGCACCAACACACAGCGCCAGCCTGGTTTA	3420
QY	2281	GCTTCTCCCCAGCTGGTATTATTCCAACTGAAGCTGTGGCATGCTTATTCAGCTGCA	2340	Db	3361	CCTGTAATACAGGGTAATGCTTAGAATGATGGCACCAACACACAGCGCCAGCCTGGTTTA	3420
Db	2281	GCTTCTCCCCAGCTGGTATTATTCCAACTGAAGCTGTGGCATGCTTATTCAGCTGCA	2340	QY	3421	GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACCGCATGCGATGTATGCAAT	3480
QY	2341	TCTCTAGCCCTGCTAGTCTCGCATCGAACAGAGCTGTATCCCTTCTAGTGAGGCTAAA	2400	Db	3421	GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACCGCATGCGATGTATGCAAT	3480
Db	2341	TCTCTAGCCCTGCTAGTCTCGCATCGAACAGAGCTGTATCCCTTCTAGTGAGGCTAAA	2400	QY	3481	CCCAATTTACCATACAAAGGAGCAAGCCCTTCTTCTTACTTTGCAATTTCCACGGGC	3540
QY	2401	GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCTCGAGGGAATAAAGAAAAATATAA	2460	Db	3481	CCCAATTTACCATACAAAGGAGCAAGCCCTTCTTCTTACTTTGCAATTTCCACGGGC	3540
Db	2401	GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCTCGAGGGAATAAAGAAAAATATAA	2460	QY	3541	TCCCTTCTCAGCAGTATGCGACCCCTTAACGCTAACCTTGCAACCCACATACTTCCACCCCT	3600
QY	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTAATATCACAGTTGTT	2520	Db	3541	TCCCTTCTCAGCAGTATGCGACCCCTTAACGCTAACCTTGCAACCCACATACTTCCACCCCT	3600
Db	2461	CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTAATATCACAGTTGTT	2520	QY	3601	CAGCCTTTCAGTACCCCTCAGCAGCAGCAAGCAACATGTTGGAAGTCATCTCTGCA	3660
QY	2521	TCTGACATAGAAACAGATTGATGATTTAAAGAAATTTAAGATGATTTTATAGTTACAG	2580				

Db 3601 CAGCCTTCAGCTACCCCACTGACAGCAGCAAGCAACATGGTGGAGTCACTCTGCA 3660  
Qy 3661 CCAGTCTCTGTTACAGACCATCAGCACCAGGCGGCCAGGCTCTCCATCTGCCAGTCCA 3720  
Db 3661 CCAGTCTCTGTTACAGACCATCAGCACCAGGCGGCCAGGCTCTCCATCTGCCAGTCCA 3720  
Qy 3721 CAGCAGCAGTCAAGCAATTTACCAACGCGGGCTTGGCGCAACTCCACCTCCATGACACT 3780  
Db 3721 CAGCAGCAGTCAAGCAATTTACCAACGCGGGCTTGGCGCAACTCCACCTCCATGACACT 3780  
Qy 3781 GCTCTCAACAGCAGTCCGACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTAG 3840  
Db 3781 GCTCTCAACAGCAGTCCGACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTAG 3840  
Qy 3841 ATCCATCTCTTCAAGTTCAGCGGGCTATACCAACCCCAACATGCCCCACATGACACT 3900  
Db 3841 ATCCATCTCTTCAAGTTCAGCGGGCTATACCAACCCCAACATGCCCCACATGACACT 3900  
Qy 3901 CAGGCTCATGTACAGTCAAGATGGTTCCTTCTCATCCAACTGCCCATGGCCATGATG 3960  
Db 3901 CAGGCTCATGTACAGTCAAGATGGTTCCTTCTCATCCAACTGCCCATGGCCATGATG 3960  
Qy 3961 CTAATCAGCAGACAGCAGCAGCAGCGGCTCCAGCGCGCTCGCTCAAAAGTGCACTACAG 4020  
Db 3961 CTAATCAGCAGACAGCAGCAGCAGCGGCTCCAGCGCGCTCGCTCAAAAGTGCACTACAG 4020  
Qy 4021 CCATTTCCAGTCTTCGACAAAGCAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAAGCC 4080  
Db 4021 CCATTTCCAGTCTTCGACAAAGCAGCGCATTTCCCTTATATGAGGCAACCTTCAGTACAAGCC 4080  
Qy 4081 CAGCAGCAGCAGTGTGAAGCTGCTGAGGAACTGAGGAACTGAGGAACTGAGGAACTGAGG 4140  
Db 4081 CAGCAGCAGCAGTGTGAAGCTGCTGAGGAACTGAGGAACTGAGGAACTGAGGAACTGAGG 4140  
Qy 4141 CTTCTACTGCTTCTACCACTGGAAGCAGCAGAACTAGAACTTTCATTTATTTGTTTTT 4200  
Db 4141 CTTCTACTGCTTCTACCACTGGAAGCAGCAGAACTAGAACTTTCATTTATTTGTTTTT 4200  
Qy 4201 TAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGTAACAGTTTCACTTGCAG 4260  
Db 4201 TAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGTAACAGTTTCACTTGCAG 4260  
Qy 4261 TGAAGATATCTGAGCAGGAGTAGGCAATTTAGGAACTTGGGGCTTATTCATAATCCA 4320  
Db 4261 TGAAGATATCTGAGCAGGAGTAGGCAATTTAGGAACTTGGGGCTTATTCATAATCCA 4320  
Qy 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAGTTATTT 4380  
Db 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAGTTATTT 4380  
Qy 4381 ATTTTTAAATACCTTTGAAAGTCATGAACACATCAGTACGAAAAAGTAACAAGAGT 4440  
Db 4381 ATTTTTAAATACCTTTGAAAGTCATGAACACATCAGTACGAAAAAGTAACAAGAGT 4440  
Qy 4441 GATTTGCTGCTATTACTCTCTAAAAAATAAAAAA 4481  
Db 4441 GATTTGCTGCTATTACTCTCTAAAAAATAAAAAA 4481

RESULT 6  
US-09-949-016-4498  
; Sequence 4498, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4498  
; LENGTH: 4484  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4498

Query Match  
Best Local Similarity 99.2%; Score 4447; DB 3; Length 4484;  
Matches 4475; Conservative 0; Mismatches 5; Indels 5; Gaps 2;

Qy 1 ACCCCCGAGAAAGCAACCCAGCGCGCCCGCTCTCAAGTGTCCCTTCCCGGCCCCCGG 60  
Db 1 ACCCCCGAGAAAGCAACCCAGCGCGCCCGCTCTCAAGTGTCCCTTCCCGGCCCCCGG 60  
Qy 61 GCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120  
Db 61 GCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120  
Qy 121 CACCTCCGCTCCACCCCGCGGCTCGCGCGCCCGCTCCGATCGCTCAGCGCGCGCA 180  
Db 121 CACCTCCGCTCCACCCCGCGGCTCGCGCGCCCGCTCCGATCGCTCAGCGCGCGCA 180  
Qy 181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTCGCCGACGCCAGGTGGCCC 240  
Db 181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTCGCCGCTTCGCCGACGCCAGGTGGCCC 239  
Qy 241 GGGTGGCGTCTGCTCCAGCGCGCGCGCGCGGAGGGCGGGCGGGCGGGTGGCGGGCG 300  
Db 241 GGGTGGCGTCTGCTCCAGCGCGCGCGCGCGGAGGGCGGGCGGGCGGGTGGCGGGCG 299  
Qy 301 CCGGAGACGTATCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
Db 301 CCGGAGACGTATCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359  
Qy 361 CAGAGTCTCCCTCCCTCCGCTCAGACTGTGTTTGGTAGCAACGGCAACGGCGCGCGCG 420  
Db 361 CAGAGTCTCCCTCCCTCCGCTCAGACTGTGTTTGGTAGCAACGGCAACGGCGCGCGCG 419  
Qy 421 TTTTGGCGCGTCTCCCGCGCGCTCTTGGTCTCGGGGGCGCTCCCGCGCGCGCTTCTGTCGTC 480  
Db 421 TTTTGGCGCGTCTCCCGCGCGCTCTTGGTCTCGGGGGCGCTCCCGCGCGCGCTTCTGTCGTC 479  
Qy 481 GTCTTCTCTCCCTCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
Db 481 GTCTTCTCTCCCTCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 539  
Qy 541 CTCGGCGCGGTGCGTCCCGCGCGCTCTCCGGCGTCTCTTGGCGCGCGCGCGCGCGCG 600  
Db 541 CTCGGCGCGGTGCGTCCCGCGCGCTCTCCGGCGTCTCTTGGCGCGCGCGCGCGCGCG 599  
Qy 601 TGTTCCTCCCGCGCGTTCGAGCGGTGATGGCGCGCTCACCATGTCTGCTGAAGCCCGAG 660  
Db 601 TGTTCCTCCCGCGCGTTCGAGCGGTGATGGCGCGCTCACCATGTCTGCTGAAGCCCGAG 659  
Qy 661 CAG 720  
Db 661 CAG 719  
Qy 721 CAGCGCGCGCGCGGTGCTCAATGTCCGAAAGCCCGCGCGCGCGCGCGCTTCTTAGCGTTCG 780  
Db 721 CAGCGCGCGCGCGGTGCTCAATGTCCGAAAGCCCGCGCGCGCGCGCTTCTTAGCGTTCG 779  
Qy 781 CCGCGCGCGCGCGCTTTCGCGCGTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCC 840  
Db 781 CCGCGCGCGCGCGCTTTCGCGCGTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCC 839  
Qy 841 TCCTCGGTGCTCGCGCGAGCTCCCGCGGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 900  
Db 841 TCCTCGGTGCTCGCGCGAGCTCCCGCGGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG

Db	840	TCCTCGTGTGCGCGGACCTCCGGCGGCGGGAGGCCCGGCTGGGCAGAGGTCGAAC	899	QY	1777	TCCTGTCTCTACTATGCTAAACCGCATGTCTTTCAGAAAGGGCTCCAAAGGATGTCCCCAA	2036
QY	901	AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATAGAGG	960	Db	1980	TCCTGTCTCTACTATGCTAAACCGCATGTCTTTCAGAAAGGGCTCCAAAGGATGTCCCCAA	2039
Db	900	AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATAGAGG	959	QY	2037	GGCCAGCGCATCCTCGAAATCACAGATTTCTGCTGGAGGGGTCCCATATCCAGTGG	2096
QY	961	ATGTTTCATATCTTACATCAGTGTGCTCCAAATGTGAAGTCAAGTGAAGAAATGGA	1020	Db	2040	GGCCAGCGCATCCTCGAAATCACAGATTTCTGCTGGAGGGGTCCCATATCCAGTGG	2099
Db	960	ATGGTTTCATATCTTACATCAGTGTGCTCCAAATGTGAAGTCAAGTGAAGAAATGGA	1019	QY	2097	CCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGCTACTCCTCCAGTAGCAAGAC	2156
QY	1021	GGTATATGAAGAGTTTTAAAACTTACAGTCCGAAGTGTGATTTGCTACTTGATGCC	1080	Db	2100	CCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGCTACTCCTCCAGTAGCAAGAC	2159
Db	1020	GGTATATGAAGAGTTTTAAAACTTACAGTCCGAAGTGTGATTTGCTACTTGATGCC	1079	QY	2157	CAGTCCCTCGGGGGAAACGTGGTCAATCAGTGGTTCAGTGGGTTCCAAGATTTATCCCTAA	2216
QY	1081	GCACATGAGAAAGTACAGAAATCCAGTTTCGGGGCCGAAACGTGAAGAAATTAATGGAGGT	1140	Db	2160	CAGTCCCTCGGGGGAAACGTGGTCAATCAGTGGTTCAGTGGGTTCCAAGATTTATCCCTAA	2219
Db	1080	GCACATGAGAAAGTACAGAAATCCAGTTTCGGGGCCGAAACGTGAAGAAATTAATGGAGGT	1139	QY	2217	AACTCATAGACCCAGGTCCTCCAGACAGAACAGTATTTGGAAATACCCCGAGTGGGCCAGT	2276
QY	1141	ATTTTGTTCAAATGTTACAGCTTTGTGTGTGATGAGTTTAAAGATATGGACTCCAGTTAT	1200	Db	2220	AACTCATAGACCCAGGTCCTCCAGACAGAACAGTATTTGGAAATACCCCGAGTGGGCCAGT	2279
Db	1140	ATTTTGTTCAAATGTTACAGCTTTGTGTGTGATGAGTTTAAAGATATGGACTCCAGTTAT	1199	QY	2277	TCTTGTCTTCTCCCAAGCTGGTATTTTCCAACTGAAAGCTGTTGCCATGCTATTTCAGC	2336
QY	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTGCTTAAAGTGAATGGCGAAACACAA	1260	Db	2280	TCTTGTCTTCTCCCAAGCTGGTATTTTCCAACTGAAAGCTGTTGCCATGCTATTTCAGC	2339
Db	1200	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTGCTTAAAGTGAATGGCGAAACACAA	1259	QY	2337	TGCATCTCTACGGCTGCTAGTCTCTGATCGAAACAGAGCTGTTTACCCCTTCTAGTGAGGC	2396
QY	1261	GAGAAAGGACTGGAGCCCTGGGATCGAGTGAATCTCACAGCCAAATGAGGAACTTGAGGCT	1320	Db	2340	TGCATCTCTACGGCTGCTAGTCTCTGATCGAAACAGAGCTGTTTACCCCTTCTAGTGAGGC	2399
Db	1260	GAGAAAGGACTGGAGCCCTGGGATCGAGTGAATCTCACAGCCAAATGAGGAACTTGAGGCT	1319	QY	2397	TAAAGATTCAGGCTTCAAGATCAGAGCGAGAACTCTCCTGCAGGGAATAAAGAAAAATAT	2456
QY	1321	TTGGAAAAATGAGTATCTAAATGGATGGGATCCCAATGATATGTTTCGATATATAATGAAG	1380	Db	2400	TAAAGATTCAGGCTTCAAGATCAGAGCGAGAACTCTCCTGCAGGGAATAAAGAAAAATAT	2459
Db	1320	TTGGAAAAATGAGTATCTAAATGGATGGGATCCCAATGATATGTTTCGATATATAATGAAG	1379	QY	2457	TAAACCCAAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACCAAAAGGTATATCACCAGT	2516
QY	1381	AAATATGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTTAGAA	1440	Db	2460	TAAACCCAAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACCAAAAGGTATATCACCAGT	2519
Db	1380	AAATATGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTTAGAA	1439	QY	2517	TGTTTCTGAAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGGTT	2576
QY	1441	AGAGATAACTCAGAAAGAAATTTTAAACCGGGAAGCAAGGGCAAAACAGTTAGCAGAAAG	1500	Db	2520	TGTTTCTGAAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGGTT	2579
Db	1440	AGAGATAACTCAGAAAGAAATTTTAAACCGGGAAGCAAGGGCAAAACAGTTAGCAGAAAG	1499	QY	2577	ACAGCCAGTTCTACTTCTGATCTATGGATCAACTACTTAACAAAAATAGAGAGGAGA	2636
QY	1501	ATTGAGTCAAGTGCCTCAGTACAAAGCTCAGTGGGCCCTGGAAATGATAGAGGTAG	1560	Db	2580	ACAGCCAGTTCTACTTCTGATCTATGGATCAACTACTTAACAAAAATAGAGAGGAGA	2639
Db	1500	ATTGAGTCAAGTGCCTCAGTACAAAGCTCAGTGGGCCCTGGAAATGATAGAGGTAG	1559	QY	2637	AAAAATCAAGAGATTGATCAAGACAAAAATTTGAACCAAGTCTTAAGGATTCCTTTCATTGA	2696
QY	1561	GAAGAAAAATACACAGCAGTTCCAGAGAAATTCACGTGAACGTGAGGGGCACAGCATAAAC	1620	Db	2640	AAAAATCAAGAGATTGATCAAGACAAAAATTTGAACCAAGTCTTAAGGATTCCTTTCATTGA	2699
Db	1560	GAAGAAAAATACACAGCAGTTCCAGAGAAATTCACGTGAACGTGAGGGGCACAGCATAAAC	1619	QY	2697	AAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCC	2756
QY	1621	ACTAGGAAAAATAATATATTCCTCTCGACAAAGAAATAGAGAAAGTCATATCCTTGGGGA	1680	Db	2700	AAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCC	2759
Db	1620	ACTAGGAAAAATAATATATTCCTCTCGACAAAGAAATAGAGAAAGTCATATCCTTGGGGA	1679	QY	2757	TTCAATCTTAGTAAACAGGAGCACAAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA	2816
QY	1681	AGTGGAGACAGAAATTCACCGGATATGGCCAGCTCGAGTGGGCTCAATGCCATCAAGA	1740	Db	2760	TTCAATCTTAGTAAACAGGAGCACAAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA	2819
Db	1680	AGTGGGAGACAGAAATTCACCGGATATGGGCCAGCTCGAGTGGGCTCCATGCCATCAAGA	1739	QY	2817	GACTTCAGCCCGCAGCATGTAAACAAAGAGAACGATTAAGGAAGAGAAAGAACGCGAGC	2876
QY	1741	TCCACTTCTCACATTTACAGATTTCAACCCGAATTCGTGTTTCAGACCAAGAGAGTAAAT	1800	Db	2820	GACTTCAGCCCGCAGCATGTAAACAAAGAGAACGATTAAGGAAGAGAAAGAACGCGAGC	2879
Db	1740	TCCACTTCTCACATTTACAGATTTCAACCCGAATTCGTGTTTCAGACCAAGAGAGTAAAT	1799	QY	2877	TGACCAAGTTAGGAAATCAACATTTGAATTCCTCAATGCAAAAGAGGTTCAACCCACGTTCTTT	2936
QY	1801	GGAG----GTGTTCCCTGGCCATCGCTTGGCCATCTCCTTCTCTCGCCCCACCTTCTCG	1856	Db	2880	TGACCAAGTTAGGAAATCAACATTTGAATTCCTCAATGCAAAAGAGGTTCAACCCACGTTCTTT	2939
Db	1800	GGAGGCAAGTGTTCCTGGCCATCGCTTGGCCATCTCCTTCTCTCGCCCCACCTTCTCG	1859	QY	2937	CTCTCAGCCAAAGCTTCTACTACCCCACTTCACTCTGGCCTCAAGCACAACCTAGGCC	2996
QY	1857	CTACAGTCAGTCCCAACTCTCTTCACCTCGGCGAGCCACCCCTACACGGCGCCCTC	1916	Db	2940	CTCTCAGCCAAAGCTTCTACTACCCCAACTTCACTCTGGCCTCAAGCACAACCTAGGCC	2999
Db	1860	CTACAGTCAGTCCCAACTCTCTTCACCTCGGCGAGCCACCCCTACACGGCGCCCTC	1919	QY	2997	ATCTATGTTGGGTCAATCAACAGCCAACTCCAGTTTATCTCAGGCTGTGTTTGTGACCC	3056
QY	1917	CAGGCCCTCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTTCTCCAGC	1976	Db	3000	ATCTATGTTGGGTCAATCAACAGCCAACTCCAGTTTATCTCAGGCTGTGTTTGTGACCC	3059
Db	1920	CAGGCCCTCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTTCTCCAGC	1979				











Db	182	GGGGGGGGGGGGTCCC	GGCGCTTGGGCGAGAGT	CGGAACAGTAGCAAAAGCATGCGCTCAG	241	
Qy	922	TTCTACGATTTCTTTT	TGATGGAATCTATG	CAAAATATCAGGATGGTT	CATATCTTACATCA	981
Db	242	CCCTACGATTTCTTTT	TGATGGAATCTATG	CAAAACGTCGAGGATGGTT	CATATCTTACGTCA	301
Qy	982	GTTTGTGGCTCCAAAT	TGTGAAGTACAAGT	GAAATAATGGAAGTATATAT	TATGAAGAGTTTTT	1041
Db	302	GTTTGTGGATCGAAAT	TGTGAAGTACAAGT	GAAATAATGGAAGTATATAT	TATGAAGAGTTTTT	361
Qy	1042	AAAACTTACAGTCCGA	AGTGTGATTTTGTA	TCTTACTTGATGCCGCA	CATGAGAAAAAGTACAGAA	1101
Db	362	AAAAACAATACAGTCT	TAAGTGTGACTTGGT	TACTTGTATGCTGCA	CATGAGAAAGTACAGAA	421
Qy	1102	TCCAGTTCGGGGCCG	AAACGCTGAAGAAAT	TAATGGAGAGTATTTT	TGTTCAAAATGTTCCAGAC	1161
Db	422	TCCAGTTCGGGGCCAA	ACGTAAGAAATAAT	GGAGAGTGTTTGT	TGTTCAAAATGCTCAGAC	481
Qy	1162	TTTGTGTGTTACAGT	TTTAAAGATATGGA	TCTCAGTTATGCAAA	AGAGATGCTTTTACT	1221
Db	482	TTCTGTGTGTTACAG	TTTAAAGATACAGAC	TCTCAGTTATGCA	CGGAGAGATGCTTTTACT	541
Qy	1222	GACTCTGCTATCAGT	GCTTAAAGTGAAT	TGGCGAAACAA	AAGAGGAAGCACTTGGAGCCCTGG	1281
Db	542	GACTCTGCTCTCAG	CGCAAGGTGAAT	TGGTGAGCA	CAAGGAGGAAGCACTTGGAGCCCTGG	601
Qy	1282	GATGCAAGTGAACT	CAAGCAATAGAGAA	CTTGAGGCTTTGGA	AAATGACGTATCTAAT	1341
Db	602	GATGCAAGGAGCT	CAAGGCGAGGAGCT	GGAGCTGGAG--	CTGGAGAAATGATGTGTCTAAT	658
Qy	1342	GGATGGGATCCCAAT	GATATATGTTTCGA	TATAATGAGAAAAAT	TATGGTGTAGTGTCTACG	1401
Db	658	GGATGGGATCCCAAT	GATATATGTTTCGA	TATAATGAGAAAAAT	TATGGTGTAGTGTCTACCA	718
Qy	1402	TATGATAGCAGTTT	ATCTTCGTATACAGT	GCCCTTAGAAA	GAGATAACTCAGAAGAATTT	1461
Db	718	TATGATAGCAGTTT	ATCTTCGTATACAGT	GCCCTTAGAAA	AGGCAACTCAGAAGAATTT	778
Qy	1462	TTAAACCGGAGCAAG	GGCAACCACTT	TAGCAGAGAAAT	TGAGTCAAGTGCCTCAGTAC	1521
Db	778	CTTAAACCGGAGCAAG	GGCAACCACTT	TAGCAGAGAAAT	TGAGTCAAGTGCCTCAGTAC	838
Qy	1522	AAAGCTCGAGTGGCC	CTTGAAAAATGAT	GATAGGAGT	CAGGAAGAAAAATACACAGCAGTT	1581
Db	838	AAAGCTCGTTCGCCC	TTGAGATGATGA	CCGGAGT	CGGAGAGAAAAATACACAGCAGTC	898
Qy	1582	CAGAGAAATTCAGT	GAAACGTGAGGGGCA	CAGCAATAACA	CTAGGGAAAAATAATATATT	1641
Db	898	CAGAGAAATTCAGT	GAAACGTGAGGGGCA	CAGCAATAACA	CTAGGGAAAAATAATATATT	958
Qy	1642	CCTCTCGACAAAGAA	TAGAGAGTCAAT	CTCTGGGGAAGT	GGGACAGAAATTCACCG	1701
Db	958	CCTCTCGACAAAGAA	TAGAGAGTCAAT	CTCTGGGGAAGT	GGGACAGAGCTCACCA	1018
Qy	1702	CGTATGGCCAGCCT	CGGATCGGGCT	TCCATGCAT	CAAGATCCATCTTCTCACACTTCCAGAT	1761
Db	1018	CGATGGCCAGCCT	GGCCAGGCTCC	ATGCCGT	CAAGAGCTGCTTCTCACACTTCCAGAT	1078
Qy	1762	TTCAACCCGGAATTC	TGGTTTCHAGACAA	AGAGTAGTTTAA	TGAGAGTGTTCCTTGGCCATCG	1821
Db	1078	TTCAACCCGGAATTC	TGGTTTCHAGACAA	AGAGTAGTTTAA	TGAGAGTGTTCCTTGGCCATCG	1138
Qy	1822	CTTTGCCCATCTCT	CTCTCTCGCCAC	CTTCTCTCGCT	TACCAAGTCAAGTCCCACTCTCTT	1881
Db	1138	CTTTGCCCATCTCT	CTCTCTCGCCAC	CTTCTCTCGCT	TACCAAGTCAAGTCCCACTCTCTT	1198
Qy	1882	CCACTTCGGGAGCC	ACCCCTTACACGG	CGCCCTCC	1917	
Db	1198	CCACTTCGGGAGCC	ACCCCTTACACGG	CGCCCTCGTGCC	1234	

```

1 / Sequence 19, Application US/08981998A
2 / Patent No. 6844431
3 / GENERAL INFORMATION:
4 / APPLICANT: PULST, STEFAN M.
5 / TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCERBELLIAR
6 / ATAXIA-2 AND PRODUCTS RELATED THERETO
7 /
8 / NUMBER OF SEQUENCES: 19
9 / CORRESPONDENCE ADDRESS:
10 / ADDRESSEE: MUETING, RAASCH & GERHARDT, P.A.
11 / STREET: 119 NORTH FOURTH STREET, SUITE 203
12 / CITY: MINNEAPOLIS
13 / STATE: MINNESOTA
14 / COUNTRY: 55401
15 / ZIP: 55401
16 / COMPUTER READABLE FORM:
17 / MEDIUM TYPE: Floppy disk
18 / COMPUTER: IBM PC compatible
19 / OPERATING SYSTEM: PC-DOS/MS-DOS
20 / SOFTWARE: PatentIn Release #1.0, Version #1.30
21 /
22 / CURRENT APPLICATION DATA:
23 / APPLICATION NUMBER: US/08/981,998A
24 / FILING DATE: 11-May-1998
25 / CLASSIFICATION: <Unknown>
26 /
27 / PRIOR APPLICATION DATA:
28 / APPLICATION NUMBER: WO 97/42314
29 / FILING DATE: 08-MAY-1997
30 / APPLICATION NUMBER: US 08/727,084
31 / FILING DATE: 08-OCT-1996
32 / APPLICATION NUMBER: US 60/022,207
33 / FILING DATE: 19-JUL-1996
34 / APPLICATION NUMBER: US 60/017,388
35 / FILING DATE: 08-MAY-1996
36 / ATTORNEY/AGENT INFORMATION:
37 / NAME: MUETING, ANN M.
38 / REGISTRATION NUMBER: 33,977
39 / REFERENCE/DOCKET NUMBER: 232.00010120
40 / TELECOMMUNICATION INFORMATION:
41 / TELEPHONE: 612-305-1217
42 / TELEFAX: 612-305-1228
43 / INFORMATION FOR SEQ ID NO: 19:
44 / SEQUENCE CHARACTERISTICS:
45 / LENGTH: 1257 base pairs
46 / TYPE: nucleic acid
47 / STRANDEDNESS: both
48 / TOPOLOGY: both
49 / MOLECULE TYPE: cDNA
50 / FEATURE:
51 / NAME/KEY: CDS
52 / LOCATION: 2..1255
53 / SEQUENCE DESCRIPTION: SEQ ID NO: 19:
54 /
55 /
56 /
57 /
58 /
59 /
60 /
61 /
62 /
63 /
64 /
65 /
66 /
67 /
68 /
69 /
70 /
71 /
72 /
73 /
74 /
75 /
76 /
77 /
78 /
79 /
80 /
81 /
82 /
83 /
84 /
85 /
86 /
87 /
88 /
89 /
90 /
91 /
92 /
93 /
94 /
95 /
96 /
97 /
98 /
99 /
100 /
101 /
102 /
103 /
104 /
105 /
106 /
107 /
108 /
109 /
110 /
111 /
112 /
113 /
114 /
115 /
116 /
117 /
118 /
119 /
120 /
121 /
122 /
123 /
124 /
125 /
126 /
127 /
128 /
129 /
130 /
131 /
132 /
133 /
134 /
135 /
136 /
137 /
138 /
139 /
140 /
141 /
142 /
143 /
144 /
145 /
146 /
147 /
148 /
149 /
150 /
151 /
152 /
153 /
154 /
155 /
156 /
157 /
158 /
159 /
160 /
161 /
162 /
163 /
164 /
165 /
166 /
167 /
168 /
169 /
170 /
171 /
172 /
173 /
174 /
175 /
176 /
177 /
178 /
179 /
180 /
181 /
182 /
183 /
184 /
185 /
186 /
187 /
188 /
189 /
190 /
191 /
192 /
193 /
194 /
195 /
196 /
197 /
198 /
199 /
200 /
201 /
202 /
203 /
204 /
205 /
206 /
207 /
208 /
209 /
210 /
211 /
212 /
213 /
214 /
215 /
216 /
217 /
218 /
219 /
220 /
221 /
222 /
223 /
224 /
225 /
226 /
227 /
228 /
229 /
230 /
231 /
232 /
233 /
234 /
235 /
236 /
237 /
238 /
239 /
240 /
241 /
242 /
243 /
244 /
245 /
246 /
247 /
248 /
249 /
250 /
251 /
252 /
253 /
254 /
255 /
256 /
257 /
258 /
259 /
260 /
261 /
262 /
263 /
264 /
265 /
266 /
267 /
268 /
269 /
270 /
271 /
272 /
273 /
274 /
275 /
276 /
277 /
278 /
279 /
280 /
281 /
282 /
283 /
284 /
285 /
286 /
287 /
288 /
289 /
290 /
291 /
292 /
293 /
294 /
295 /
296 /
297 /
298 /
299 /
300 /
301 /
302 /
303 /
304 /
305 /
306 /
307 /
308 /
309 /
310 /
311 /
312 /
313 /
314 /
315 /
316 /
317 /
318 /
319 /
320 /
321 /
322 /
323 /
324 /
325 /
326 /
327 /
328 /
329 /
330 /
331 /
332 /
333 /
334 /
335 /
336 /
337 /
338 /
339 /
340 /
341 /
342 /
343 /
344 /
345 /
346 /
347 /
348 /
349 /
350 /
351 /
352 /
353 /
354 /
355 /
356 /
357 /
358 /
359 /
360 /
361 /
362 /
363 /
364 /
365 /
366 /
367 /
368 /
369 /
370 /
371 /
372 /
373 /
374 /
375 /
376 /
377 /
378 /
379 /
380 /
381 /
382 /
383 /
384 /
385 /
386 /
387 /
388 /
389 /
390 /
391 /
392 /
393 /
394 /
395 /
396 /
397 /
398 /
399 /
400 /
401 /
402 /
403 /
404 /
405 /
406 /
407 /
408 /
409 /
410 /
411 /
412 /
413 /
414 /
415 /
416 /
417 /
418 /
419 /
420 /
421 /
422 /
423 /
424 /
425 /
426 /
427 /
428 /
429 /
430 /
431 /
432 /
433 /
434 /
435 /
436 /
437 /
438 /
439 /
440 /
441 /
442 /
443 /
444 /
445 /
446 /
447 /
448 /
449 /
450 /
451 /
452 /
453 /
454 /
455 /
456 /
457 /
458 /
459 /
460 /
461 /
462 /
463 /
464 /
465 /
466 /
467 /
468 /
469 /
470 /
471 /
472 /
473 /
474 /
475 /
476 /
477 /
478 /
479 /
480 /
481 /
482 /
483 /
484 /
485 /
486 /
487 /
488 /
489 /
490 /
491 /
492 /
493 /
494 /
495 /
496 /
497 /
498 /
499 /
500 /
501 /
502 /
503 /
504 /
505 /
506 /
507 /
508 /
509 /
510 /
511 /
512 /
513 /
514 /
515 /
516 /
517 /
518 /
519 /
520 /
521 /
522 /
523 /
524 /
525 /
526 /
527 /
528 /
529 /
530 /
531 /
532 /
533 /
534 /
535 /
536 /
537 /
538 /
539 /
540 /
541 /
542 /
543 /
544 /
545 /
546 /
547 /
548 /
549 /
550 /
551 /
552 /
553 /
554 /
555 /
556 /
557 /
558 /
559 /
560 /
561 /
562 /
563 /
564 /
565 /
566 /
567 /
568 /
569 /
570 /
571 /
572 /
573 /
574 /
575 /
576 /
577 /
578 /
579 /
580 /
581 /
582 /
583 /
584 /
585 /
586 /
587 /
588 /
589 /
590 /
591 /
592 /
593 /
594 /
595 /
596 /
597 /
598 /
599 /
600 /
601 /
602 /
603 /
604 /
605 /
606 /
607 /
608 /
609 /
610 /
611 /
612 /
613 /
614 /
615 /
616 /
617 /
618 /
619 /
620 /
621 /
622 /
623 /
624 /
625 /
626 /
627 /
628 /
629 /
630 /
631 /
632 /
633 /
634 /
635 /
636 /
637 /
638 /
639 /
640 /
641 /
642 /
643 /
644 /
645 /
646 /
647 /
648 /
649 /
650 /
651 /
652 /
653 /
654 /
655 /
656 /
657 /
658 /
659 /
660 /
661 /
662 /
663 /
664 /
665 /
666 /
667 /
668 /
669 /
670 /
671 /
672 /
673 /
674 /
675 /
676 /
677 /
678 /
679 /
680 /
681 /
682 /
683 /
684 /
685 /
686 /
687 /
688 /
689 /
690 /
691 /
692 /
693 /
694 /
695 /
696 /
697 /
698 /
699 /
700 /
701 /
702 /
703 /
704 /
705 /
706 /
707 /
708 /
709 /
710 /
711 /
712 /
713 /
714 /
715 /
716 /
717 /
718 /
719 /
720 /
721 /
722 /
723 /
724 /
725 /
726 /
727 /
728 /
729 /
730 /
731 /
732 /
733 /
734 /
735 /
73
```

```
|||||
242 CCTACGATTTCTTTTATGAGAACTATGCAACGCTGAGATGGTTCATATATCTTACGTCA 301
QY 982 GTTGTGGCTCAAAATGTCAAGTACAAGTGAATAATGAGAGTATATATCAAGAGTCTTTT 1041
Db 302 GTTGTGGATCGAAATGTCAAGTACAAGTGAATAATGAGAGTATATATGAAGAGTCTTTT 361
QY 1042 AAAACTTACAGTCCGAAGTGTGATTTGGTACTTGTATGCGGCACATGAGAAAAGTACAGAA 1101
Db 362 AAAACATACAGTCCTAAGTGTGACTTGTGTAATGCTGACATGAGAAAAGTACAGAA 421
QY 1102 TCCAGTTCGGGGCCGAAAGCTGAAGAAATAATGAGAGTATTTTCTCAAAATGTTTCAGAC 1161
Db 422 TCCAGTTCGGGGCCGAAAGCTGAAGAAATAATGAGAGTGTTTTCTCAAAATGCTCAGAC 481
QY 1162 TTTGTGTGGTACAGTTTAAAGATATGGAATCCAGTTCAGTATGCAAAAAGAGATGCTTTTACT 1221
Db 482 TTTGTGTGGTACAGTTTAAAGATACAGACTCCAGTTATGACGGAGAGATGCTTTTACT 541
QY 1222 GACTCTGCTATCAGTGTCTAAAGTGAATGGCGAACAACAAGAGAGGACCTGGAGCCCTGG 1281
Db 542 GACTCTGCTCTCAGCGCAAAAGTGAATGGTGAGCAACAAGGAGAGGACCTGGAGCCCTGG 601
QY 1282 GATGCAAGTGAATCACAGCCCAATGAGGAACCTTGAGGCTTTGGAAATGACGTATCTAAT 1341
Db 602 GATGCAAGGAGCTCACGGCCAGCGAGAGCTGGAG---CTGGAGAAATGATGTGCTAAT 658
QY 1342 GGATGGGATCCCAATGATATGTTTCGATATATGAAGAAAATATGCTGTAGTGTCTACG 1401
Db 659 GGATGGGACCCCAATGACATGTTTCGATATATGAAGAAATATGCTGTGTGCTCCACA 718
QY 1402 TATGATAGAGTTTATCTTCGATATACAGTGCCTTTAGAAAAGAGATAACTCAGAGAAATTT 1461
Db 719 TATGATAGAGTTTATCTTCATATACGTTCTTTAGAAAAGGACAACTCAGAGAAATTT 778
QY 1462 TTAAACCGGAGCAAGGGCAACACAGTTAGCAGAAAGAAATGAGTCAAGTCCCGAGTAC 1521
Db 779 CTTAACCGGGAGGCAAGGGCAACACAGTTAGCAGAAAGAAATGAAATCCAGTGTCTCAGTAC 838
QY 1522 AAAGCTCGAGTGGCCCTCGAAAATGATAGAGAGTGAAGAAAGAAAATACACAGCAGTT 1581
Db 839 AAAGCTCGTGTGCGCCTTGAGATGATGACCGGAGTGAAGAAAGAAAATACACAGCAGTC 898
QY 1582 CAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAACACTAGGGAATAATAATATATT 1641
Db 899 CAGAGAAATCTCAGTGAACCGGGAGGGCATGCCCCCAACACTAGGCACAATAATAATATT 958
QY 1642 CCTCTGGACAAAGAAATAGAGAACTCATCTCGGGAGTGGAGAGACAGAAATTCACCG 1701
Db 959 CCTCTGGACAAAGAAACAGAGAAATGCTCTATCTCGGGAAAGTGGGAGACAGAGCTCACCA 1018
QY 1702 CGTATGGGCGAGCTGGATCGGGCTCCATGCGCATCAAGATCCACTTCTCACACTTTCAGAT 1761
Db 1019 CGGATGGGCGAGCTTGGGCCAGGCTCCATGCGCTCAGAGCTGCTTCTCACACTTTCAGAT 1078
QY 1762 TTCAACCCGAAATTCGTTTTCAGACCAAGAGTGAATGAGAGTGTTCCTTGGGCCATCG 1821
Db 1079 TTCAACCCGAAACGCTGGGCTCAGACCAAGAGTGAATGAGAGTGTTCCTTGGGCCATCG 1138
QY 1822 CCTTGGCCATCTCTTCTCTCGGCCCATCTTCTCGCTACCACTCAGTCCCACTCTCTT 1881
Db 1139 CCTTGGCCATCTCTTCTCTCGGCCCATCTTCTCGCTACCACTCAGTCCCACTCTCTT 1198
QY 1882 CCACCTCGGGCAGGCCACCCCTACAGCGCGCCCTCC 1917
Db 1199 CCACCTCGGGCAGGCCACCCCTACAGCGGCTCGTGCC 1234
```

## RESULT 11

US-09-949-016-16240

; Sequence 16240, Application US/09949016

; Patent No. 681239

; GENERAL INFORMATION:

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16240
; LENGTH: 151088
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(151088)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16240
```

Query Match 19.4%; Score 870; DB 3; Length 151088;

Best Local Similarity 99.3%; Pred. No. 1.1e-179;

Matches 884; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
QY 1 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCTCCAGTGCCTCCAGTGCCTCCCGGCCCGGG 60
Db 2001 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCTCCAGTGCCTCCCGGCCCGGG 2060
QY 61 GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
Db 2061 GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 2120
QY 121 CACTCTCGCTCCCAACCGCGGCCCTTCGGCGGCCCGCCCTCCGATGCGTCAAGCGGCCGA 180
Db 2121 CACTCTCGCTCCCAACCGCGGCCCTTCGGCGGCCCGCCCTCCGATGCGTCAAGCGGCCGA 2180
QY 181 GCTCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGGACGAGCGGCGCC 240
Db 2181 GCTCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGG-TTCGCGGACGAGCGGCGCC 2239
QY 241 GGGTGGGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCG 300
Db 2240 GGGTGGGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 2299
QY 301 CCGGAGCGTATCCCTTCGCGGCCCTTCCTCCCGCCCGGCCCGCCCTTCCCTTCCCGG 360
Db 2300 CCGGAGCGGATCCCTTCGCGGCCCTTCCTCCCGCCCGGCCCGCCCTTCCCTTCCCGG 2359
QY 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCAAGCGCGGCGGCGG 420
Db 2360 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCAAGCGCGGCGGCGG 2419
QY 421 TTTCCGCGCGGCTCCCGCGGCTCTTGTGCTCGCGGCGGCTCCCGGCCCTTCCGTGCTC 480
Db 2420 TTTCCGCGCGGCTCCCGCGGCTCTTGTGCTCGCGGCGGCTCCCGGCCCTTCCGTGCTC 2479
QY 481 GTCCTTCTCCCTTCGCGAGCCCGGGCGCCCTTCGGCGGCGGCGCAACCGCGGCTTCCCG 540
Db 2480 CTCCTTCTCCCTTCGCGAGCCCGGGCGCCCTTCGGCGGCGGCGCAACCGCGGCTTCCCG 2539
QY 541 CTCGGCGCGGTGCTCCCGCGGCTTCGGCGGCTTCTTGGCGCGCCCGGCTCCCGG 600
Db 2540 CTCGGCGCGCGGCTTCGCGGCTTCGGCGGCTTCTTGGCGCGCCCGGCTCCCGG 2599
QY 601 TGTCCCGCGCGGCTGCGAGCGGTGATGGGCGCCCTCACCATGTGCTGAAGCCCCAG 660
Db 2600 TGTCCCGCGCGGCTGCGAGCGGTGATGGGCGCCCTCACCATGTGCTGAAGCCCCAG 2659
QY 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
```





```
QY 375 CTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGGGGGGGGTTTCGGCCGGGCTC 434
Db 1 CTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGGGGGGGGTTTCGGCCGGGCTC 60
QY 435 CCGGGGGCTCCTTGGTCTCGGGGGGCTCCCGGCCCTTCGTGCTCGTCTCTTCTCCCT 494
Db 61 CCGGGGGCTCCTTGGTCTCGGGGGGCTCCCGGCCCTTCGTGCTCGTCTCTTCTCCCT 120
QY 495 CGCCAGCCCGGGGGCCCTTCGGCCCGCGCAACCGCGGCTCCCGCTCGGGCCGGTGC 554
Db 121 CGCCAGCCCGGGGGCCCTTCGGCCCGCGCAACCGCGGCTCCCGCTCGGGCCGGC 180
QY 555 GTCCCCCGCGGTTCCGGGCTCTCCTTGGGGGGCCCGGCTCCCGGCTGTCGGCCCGGC 614
Db 181 GTCCCCCGCGGTTCCGGGCTCTCCTTGGGGGGCCCGGCTCCCGGCTGTCGGCCCGGC 240
QY 615 GTCCGAGCCGGTGATGGGCCCTCACCATGTGCTGAAGCCCGAGCAGCAGCAGCA 674
Db 241 GTCCGAGCCGGTGATGGGGCCCTCACCATGTGCTGAAGCCCGAGCAGCAGCAGCA 300
QY 675 GCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 734
Db 301 GCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360
QY 735 GGCTGCCAATGTCGCAAGCCCGGGCGGAGCGGCTTCTAGGTCGGCCCGCGCGCGCC 794
Db 361 GGCTGCCAATGTCGCAAGCCCGGGCGGAGCGGCTTCTAGGTCGGCCCGCGCGCGCC 420
QY 795 TTCGCGTCTCTGCTCGGCTCTCCTCGTCTCGTCTCGGCGAC 833
Db 421 TTCGCGTCTCTGCTCGGCTCTCCTCGTCTCGTCTCGGCGAC 459
```

Search completed: December 23, 2005, 05:15:14  
Job time : 784 secs

**This Page Blank (uspto)**



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	4481	100.0	4481	6	US-10-141-541-5		Sequence 5, Appli
2	4481	100.0	4481	8	US-10-750-323-2		Sequence 2, Appli
3	4481	100.0	4481	9	US-10-887-553A-493		Sequence 493, App
4	4481	100.0	4481	9	US-10-802-228-1		Sequence 1, Appli
5	3138.2	70.0	4425	6	US-10-141-541-3		Sequence 3, Appli
6	1730.6	38.6	2006	3	US-09-925-298-15		Sequence 15, Appl
7	1730.6	38.6	2006	5	US-10-102-806-15		Sequence 15, Appl
8	935.2	20.9	1257	6	US-10-141-541-1		Sequence 1, Appli
9	935.2	20.9	1257	8	US-10-750-323-4		Sequence 4, Appli
10	497.4	11.1	516	8	US-10-750-323-1		Sequence 1, Appli
11	452	10.1	682	9	US-10-956-157-3976		Sequence 3976, Ap
12	452	10.1	682	9	US-10-956-157-9211		Sequence 9211, Ap
13	441.2	9.8	704	8	US-10-363-345A-23185		Sequence 23185, A
14	441.2	9.8	704	8	US-10-363-4A5A-23186		Sequence 23186, A
15	441.2	9.8	704	9	US-10-363-483A-23185		Sequence 23185, A
16	441.2	9.8	704	9	US-10-363-483A-23186		Sequence 23186, A
17	437.6	9.8	485	3	US-09-918-995-28268		Sequence 28268, A
18	412.4	9.2	73995	5	US-10-087-192-208		Sequence 208, App
19	381	8.5	465	4	US-09-925-065A-129085		Sequence 129085
20	301.4	6.7	305	7	US-10-242-535A-34580		Sequence 34580, A
21	301.4	6.7	305	7	US-10-085-783A-34580		Sequence 34580, A
22	288.2	6.4	493631	5	US-10-087-192-205		Sequence 205, App
23	281.4	6.3	2285	4	US-09-925-065A-129084		Sequence 129084

[illegible]



ADDRESSEE: Muetling, Raasch & Gebhardt, P.A.

STREET: 119 North Fourth Street

CITY: Minneapolis

STATE: Minnesota

COUNTRY: USA

ZIP: 55401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/750,323

FILING DATE: 30-Dec-2003

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/727,084

FILING DATE: 08-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Muetling, Ann M.

REGISTRATION NUMBER: 33,977

REFERENCE/DOCKET NUMBER: 232.00010101

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612/305-1220

TELEFAX: 612/305-1228

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 4481 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 163..4101

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-750-323-2

Query Match 100.0%; Score 4481; DB 8; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCCTCAGTGTCCCTCCCGCGCCCGGG	60
DB	1	ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCCTCAGTGTCCCTCCCGCGCCCGGG	60
QY	61	GCACCTCAGCTTCGTCTTCGCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG	120
DB	61	GCACCTCAGCTTCGTCTTCGCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG	120
QY	121	CACCTCCGCTCCACCCCGCGCTCGCGCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA	180
DB	121	CACCTCCGCTCCACCCCGCGCTCGCGCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA	180
QY	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTTCGCGTTCGCGTAAAGAGTCCCTATCCG	240
DB	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTTCGCGTTCGCGTAAAGAGTCCCTATCCG	240
QY	241	GGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	300
DB	241	GGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	300
QY	301	CGGGACCGTATCCCTCCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	360
DB	301	CGGGACCGTATCCCTCCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	360
QY	361	CAGACTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAGCGGCAACCGCGCGCGCGCG	420
DB	361	CAGACTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAGCGGCAACCGCGCGCGCGCG	420
QY	421	TTTGGCGCGCGCTCCCGCGCGCTTCGTCTCGCGCGCGCGCTCCCGCGCGCGCTTCGT	480
DB	421	TTTGGCGCGCGCTCCCGCGCGCTTCGTCTCGCGCGCGCGCTCCCGCGCGCGCTTCGT	480

QY	481	GTCTCTTCTCCCTCGCGCAGCCCGGGCGCCCTCCTCGGCGCGCGCAACCCCGCGCTCCCG	540
DB	481	GTCTCTTCTCCCTCGCGCAGCCCGGGCGCCCTCCTCGGCGCGCGCAACCCCGCGCTCCCG	540
QY	541	CTCGGCGCGCTGGTCCCGCGCGCTTCGGGCGCTCTCTTGGGCGCGCGCGCTCCCGCG	600
DB	541	CTCGGCGCGCTGGTCCCGCGCGCTTCGGGCGCTCTCTTGGGCGCGCGCGCTCCCGCG	600
QY	601	TGTCTCCCGCGCGCTGGAGCGGTGTATGGGCGCCCTCACCATGTCTCGCTCAAGCGCCAG	660
DB	601	TGTCTCCCGCGCGCTGGAGCGGTGTATGGGCGCCCTCACCATGTCTCGCTCAAGCGCCAG	660
QY	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG	720
DB	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG	720
QY	721	CAGCGCGCGCGCGCTGCCAATGTCCGCAAGCCCGCGCGCGCGCGCTTCTAGCGTCG	780
DB	721	CAGCGCGCGCGCGCTGCCAATGTCCGCAAGCCCGCGCGCGCGCGCTTCTAGCGTCG	780
QY	781	CCCGCGCGCGCTTCGCGCTCTCGCTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	840
DB	781	CCCGCGCGCGCTTCGCGCTCTCGCTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	840
QY	841	TCCTCGTGTGGTGGCGGACCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	900
DB	841	TCCTCGTGTGGTGGCGGACCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	900
QY	901	AGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGTATGGAATCTATGCAAAATATG	960
DB	901	AGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGTATGGAATCTATGCAAAATATG	960
QY	961	ATGTTTCATATACATCAGTTTGTGGCTCCAAATGTGAAGTACAAATGCAAAATATGGA	1020
DB	961	ATGTTTCATATACATCAGTTTGTGGCTCCAAATGTGAAGTACAAATGCAAAATATGGA	1020
QY	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGGTATCTTGTATGCC	1080
DB	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGGTATCTTGTATGCC	1080
QY	1081	GCATGAGAAAGTACAGATCCAGTTCCGGGCGGAAACGTGAAGAAATATGGAGAGT	1140
DB	1081	GCATGAGAAAGTACAGATCCAGTTCCGGGCGGAAACGTGAAGAAATATGGAGAGT	1140
QY	1141	ATTTTGTTCATATGTTTGT	1200
DB	1141	ATTTTGTTCATATGTTTGT	1200
QY	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCAAGTGAATGGCGAACAACAA	1260
DB	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCAAGTGAATGGCGAACAACAA	1260
QY	1261	GAGAGGACCTGGAGCCCTGGGATGCGAGTGAATCTCAGCAGCAATGAGGAACTTGGAG	1320
DB	1261	GAGAGGACCTGGAGCCCTGGGATGCGAGTGAATCTCAGCAGCAATGAGGAACTTGGAG	1320
QY	1321	TTGAAATATGAGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
DB	1321	TTGAAATATGAGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
QY	1381	AATTATGGTGTAGTCTTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA	1440
DB	1381	AATTATGGTGTAGTCTTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA	1440
QY	1441	AGAGATAACTCAGAAAGATTTTAAACCGGAGCAAGGCGCAACCCAGTTAGCAGAAAG	1500
DB	1441	AGAGATAACTCAGAAAGATTTTAAACCGGAGCAAGGCGCAACCCAGTTAGCAGAAAG	1500
QY	1501	ATTGAGTCAAGTGCAGTCAAAAGCTGAGTGGCGCTTGGAAATGATGATAGGAGTGA	1560
DB	1501	ATTGAGTCAAGTGCAGTCAAAAGCTGAGTGGCGCTTGGAAATGATGATAGGAGTGA	1560



Db 3721 CAGCAGCAGTACGCCATTTTACCACGGGGGCTTGCGCCAACTCCACCCCTCCATGACACCT 3780  
Qy 3781 GCCTCCAAACACGAGTCGCGCACAGAAATAGTTTCCAGCAGCAGCAACACAGACTGCTTTTACG 3840  
Db 3781 GCCTCCAAACACGAGTCGCGCACAGAAATAGTTTCCAGCAGCAGCAACACAGACTGCTTTTACG 3840  
Qy 3841 ATCCATCCTTCTCAGTTTCAGCGGGGGATATACAAACCCACATGCGCCCACTGCTTACG 3900  
Db 3841 ATCCATCCTTCTCAGTTTCAGCGGGGGATATACAAACCCACATGCGCCCACTGCTTACG 3900  
Qy 3901 CAGGCTCATGTACAGTCAGGAATGTTCTTCTCATCCAACTGCCCATGCGGCAATGATG 3960  
Db 3901 CAGGCTCATGTACAGTCAGGAATGTTCTTCTCATCCAACTGCCCATGCGGCAATGATG 3960  
Qy 3961 CTAATGAGGACACAGCCACCCGGCGGTCCCAAGCGCGCCCTCGCTCAAAGTGCACTACAG 4020  
Db 3961 CTAATGAGGACACAGCCACCCGGCGGTCCCAAGCGCGCCCTCGCTCAAAGTGCACTACAG 4020  
Qy 4021 CCATTTCCAGTCTCAGCAACAGCGGATTTCCCTTATATGACGACCCCTTCAGTACAAGCC 4080  
Db 4021 CCATTTCCAGTCTCAGCAACAGCGGATTTCCCTTATATGACGACCCCTTCAGTACAAGCC 4080  
Qy 4081 CACCACCAACAGCAGTTGTAAAGGCTGCCCTGGAGAACCGGAAAGGCCAAATTTCCCTCCTC 4140  
Db 4081 CACCACCAACAGCAGTTGTAAAGGCTGCCCTGGAGAACCGGAAAGGCCAAATTTCCCTCCTC 4140  
Qy 4141 CTTCTACTGCTTCTACCAACTGGAAGCAGCAGAAACTAGAAATTTCAITTTTGTGTTTT 4200  
Db 4141 CTTCTACTGCTTCTACCAACTGGAAGCAGCAGAAACTAGAAATTTCAITTTTGTGTTTT 4200  
Qy 4201 TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAACAGTTCACTTGCA 4260  
Db 4201 TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAACAGTTCACTTGCA 4260  
Qy 4261 TGAAGATACCTTGGACCGAGTAGGCACTTTAGGAACCTTGGGGGCTATTCCATAATTTCCA 4320  
Db 4261 TGAAGATACCTTGGACCGAGTAGGCACTTTAGGAACCTTGGGGGCTATTCCATAATTTCCA 4320  
Qy 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCCGAAACTGGAAGTTATTT 4380  
Db 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCCGAAACTGGAAGTTATTT 4380  
Qy 4381 ATTTTAAATACCTTGAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAGAGT 4440  
Db 4381 ATTTTAAATACCTTGAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAGAGT 4440  
Qy 4441 GATTCCTGCTGCTATTACTGCTTAAAAAATAAAAAAAAAAAAA 4481  
Db 4441 GATTCCTGCTGCTATTACTGCTTAAAAAATAAAAAAAAAAAAA 4481

RESULT 3  
US-10-887-553A-493  
; Sequence 493, Application US/10887553A  
; Publication No. US20050085436A1  
; GENERAL INFORMATION:  
; APPLICANT: Garza, Dan  
; APPLICANT: Li, Hao  
; TITLE OF INVENTION: Method to treat conditions associated  
; with insulin signalling dysregulation  
; FILE REFERENCE: 4-33262  
; CURRENT APPLICATION NUMBER: US/10/887,553A  
; CURRENT FILING DATE: 2004-07-08  
; PRIOR APPLICATION NUMBER: 60/485,883  
; NUMBER OF SEQ ID NOS: 1208  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 493  
; LENGTH: 4481  
; TYPE: DNA  
; ORGANISM: human  
US-10-887-553A-493

Query Match 100.0%; Score 4481; DB 9; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCCCTCCGGCCCCCGG 60  
Db 1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCCCTCCGGCCCCCGG 60  
Qy 61 GCCACCTCAGCTTCTGCTTCCGCTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
Db 61 GCCACCTCAGCTTCTGCTTCCGCTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120  
Qy 121 CACCTCCGCTCCCAACCCGGCGCCTCGGCGCGCCCTCCGATGCGGTCAAGCGCGCA 180  
Db 121 CACCTCCGCTCCCAACCCGGCGCCTCGGCGCGCCCTCCGATGCGGTCAAGCGCGCA 180  
Qy 181 GCTCCTCGGAGTCCCGGCTGGCCACCGAGTCTTCGCGCTTCGCGCAGCAGGTGGGCC 240  
Db 181 GCTCCTCGGAGTCCCGGCTGGCCACCGAGTCTTCGCGCTTCGCGCAGCAGGTGGGCC 240  
Qy 241 GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 300  
Db 241 GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGG 300  
Qy 301 CCGGGAACGATATCCCTTCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCG 360  
Db 301 CCGGGAACGATATCCCTTCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCG 360  
Qy 361 CAGAGCTCGCCTCCGCTCAGACTGTTTGGTAGCAACGCGGCGGCGGCGGCGG 420  
Db 361 CAGAGCTCGCCTCCGCTCAGACTGTTTGGTAGCAACGCGGCGGCGGCGGCGG 420  
Qy 421 TTTCCGCGCGGCTCCCGCGGCTCCTTGGTCTCGCGCGGCTCCCGCGCGCTTCGTCGTC 480  
Db 421 TTTCCGCGCGGCTCCCGCGGCTCCTTGGTCTCGCGCGGCTCCCGCGCGCTTCGTCGTC 480  
Qy 481 GTCCTTCTCCCTTCGCGCAGCCCGGCGCGCTCCCGCGCGCGCGCGCGCGCTTCCTCCG 540  
Db 481 GTCCTTCTCCCTTCGCGCAGCCCGGCGCGCTCCCGCGCGCGCGCGCGCGCTTCCTCCG 540  
Qy 541 CTCGGCGCGGTGGTTCGCGCGCGCTTCGCGCGCTCTTCGCGCGCGCGCGCGCTTCGCGG 600  
Db 541 CTCGGCGCGGTGGTTCGCGCGCGCTTCGCGCGCTCTTCGCGCGCGCGCGCGCTTCGCGG 600  
Qy 601 TGTCCCGCGCGGCTCGAGCCGCTGTATGGGCGCTTCACCATGTCGCTGAAGCGCCAG 660  
Db 601 TGTCCCGCGCGGCTCGAGCCGCTGTATGGGCGCTTCACCATGTCGCTGAAGCGCCAG 660  
Qy 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720  
Db 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720  
Qy 721 CAGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 780  
Db 721 CAGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 780  
Qy 781 CCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 840  
Db 781 CCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 840  
Qy 841 TCCTCGGTGTCGCGGACCTCCGCGCGGCGGAGCGCGCGCTGGGCGAGGTCGAAAC 900  
Db 841 TCCTCGGTGTCGCGGACCTCCGCGCGGCGGAGCGCGCGCTGGGCGAGGTCGAAAC 900  
Qy 901 AGTAACAAAGACTGCTCAGTCTACGATTTCTTTGATGAACTATGCAAAATATGAGG 960  
Db 901 AGTAACAAAGACTGCTCAGTCTACGATTTCTTTGATGAACTATGCAAAATATGAGG 960  
Qy 961 ATGTTTCATATACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTCAAAATATGGA 1020  
Db 961 ATGTTTCATATACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTCAAAATATGGA 1020  
Qy 1021 GGTATATGAAGGAGTGTATTAATAAACTTACAGTCCGGAAGTGTGTTGGTACTTGTATGCC 1080

1021	Db		GGTATATATCAAGGAGCTTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTTGTATGCC	1080
1081	Qy		GCACATGAGAAAAGTACGAATCCAGTTCGGGGCCGAAACGTTGAAGAATAATATGAGAGT	1140
1081	Db		GCACATGAGAAAAGTACGAATCCAGTTCGGGGCCGAAACGTTGAAGAATAATATGAGAGT	1140
1141	Qy		ATTTTGTTCAAAATGTTTCAGACTTTGTGTGTGACAGTTTAAAGATATGACATCCAGTTAT	1200
1141	Db		ATTTTGTTCAAAATGTTTCAGACTTTGTGTGTGACAGTTTAAAGATATGACATCCAGTTAT	1200
1201	Qy		GCAAAAAGAGATGCTTTTACTGTACTCTGTATACAGTGTCTAAAGTGAATGGCGAAACACAAA	1260
1201	Db		GCAAAAAGAGATGCTTTTACTGTACTCTGTATACAGTGTCTAAAGTGAATGGCGAAACACAAA	1260
1261	Qy		GAGAAAGACCTGAGCCCTGGGATGCAAGTGAACCTCACAGCCAAATGAGAGAACTTGAGGCT	1320
1261	Db		GAGAAAGACCTGAGCCCTGGGATGCAAGTGAACCTCACAGCCAAATGAGAGAACTTGAGGCT	1320
1321	Qy		TTGGAAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATTAATGAAGAA	1380
1321	Db		TTGGAAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATTAATGAAGAA	1380
1381	Qy		AATATATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTTAGAA	1440
1381	Db		AATATATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTTAGAA	1440
1441	Qy		AGAGATTAACCTCAGNAGNAATTTTTTAAACGGGAGCAAGGGCAAAACAGTTACGAGAAGAA	1500
1441	Db		AGAGATTAACCTCAGNAGNAATTTTTTAAACGGGAGCAAGGGCAAAACAGTTACGAGAAGAA	1500
1501	Qy		ATTGAGTCAAGTGCCTCAGTACAAAGCTCCAGTTCGAGTGGCCCTGGAAAATGATAGAGAGTAG	1560
1501	Db		ATTGAGTCAAGTGCCTCAGTACAAAGCTCCAGTTCGAGTGGCCCTGGAAAATGATAGAGAGTAG	1560
1561	Qy		GAAGAAAAATACACAGCAGTTTCAGAGAAAATTCAGGTGAACGTTGAGGGGCACAGCATAAAC	1620
1561	Db		GAAGAAAAATACACAGCAGTTTCAGAGAAAATTCAGGTGAACGTTGAGGGGCACAGCATAAAC	1620
1621	Qy		ACTAGGGAAAAATAATATATCTCTCGGACAAAGAAATAGAGAAGTCATATCCTCGGGG	1680
1621	Db		ACTAGGGAAAAATAATATATCTCTCGGACAAAGAAATAGAGAAGTCATATCCTCGGGG	1680
1681	Qy		AGTGGGAGACAGAAATTCACCGCGTATGGGCCAGCCTGGATCGGGCTCCATGCCATCAAGA	1740
1681	Db		AGTGGGAGACAGAAATTCACCGCGTATGGGCCAGCCTGGATCGGGCTCCATGCCATCAAGA	1740
1741	Qy		TCCAATTCTCACACTTCAGATTTCAACCCCGAATTCGGTTTCAGACCAAGAGTAGTTAAT	1800
1741	Db		TCCAATTCTCACACTTCAGATTTCAACCCCGAATTCGGTTTCAGACCAAGAGTAGTTAAT	1800
1801	Qy		GGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCTTCCTCTCGCCCAACCTTCTCCCTAC	1860
1801	Db		GGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCTTCCTCTCGCCCAACCTTCTCCCTAC	1860
1861	Qy		CAGTCAGTCCCAACTCTCTTCCACCTCGGGCAGCCACCCTACAGCGGCGCCCTCCAGG	1920
1861	Db		CAGTCAGTCCCAACTCTCTTCCACCTCGGGCAGCCACCCTACAGCGGCGCCCTCCAGG	1920
1921	Qy		CCCCCTCGCGGCCATCCAGACCCCCGCTCTCACCCCTCTGTCTCATGGTTTCTTCAGATCTCT	1980
1921	Db		CCCCCTCGCGGCCATCCAGACCCCCGCTCTCACCCCTCTGTCTCATGGTTTCTTCAGATCTCT	1980
1981	Qy		GTCTCTACTATGCTTAAACGATGTCTTCAGAAAGGGCTTCCAGGATGTCCCCAAAGGCC	2040
1981	Db		GTCTCTACTATGCTTAAACGATGTCTTCAGAAAGGGCTTCCAGGATGTCCCCAAAGGCC	2040
2041	Qy		CAGCGACATCTCGAAATTCACAGAGTTTCTGCTGGGAGGGGTTCCATATCCAGTGGGCTTA	2100
2041	Db		CAGCGACATCTCGAAATTCACAGAGTTTCTGCTGGGAGGGGTTCCATATCCAGTGGGCTTA	2100
2101	Qy		GAATTTGTATCCCCAACCCCAACCCAGTGAAGCAGCTACTCTCTCAGTAGCAAGCACCAAGT	2160

Db	2101	GAATTTGTATCCCAACAAACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGGACCAGT	2160
Qy	2161	CCCTCGGGGGGAAACGTGGTTCATCAGTGGTTCAGTGGGGTTTCCAAGATTATCCCTTAAAACT	2220
Db	2161	CCCTCGGGGGGAAACGTGGTTCATCAGTGGTTCAGTGGGGTTTCCAAGATTATCCCTTAAAACT	2220
Qy	2221	CATAGACCCAGGTCCTCCAGACAGAACAGATTGTGNAATACCCCCAGTGGGCGAGTTCCT	2280
Db	2221	CATAGACCCAGGTCCTCCAGACAGAACAGATTGTGNAATACCCCCAGTGGGCGAGTTCCT	2280
Qy	2281	GCCTTCTCCCAAGCTGTATTATTTCCAACTGAAGCTGTTGCCATGCCTATTCCAGCTGCA	2340
Db	2281	GCCTTCTCCCAAGCTGTATTATTTCCAACTGAAGCTGTTGCCATGCCTATTCCAGCTGCA	2340
Qy	2341	TCTCCTACGCTGCTAGTCTCTGCATCGAATCAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA	2400
Db	2341	TCTCCTACGCTGCTAGTCTCTGCATCGAATCAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA	2400
Qy	2401	GATTCCAGGCTTCAAAGTACAGAGCGAGAATCTCTCTGCAGGGAATAAAGAAAATATTAA	2460
Db	2401	GATTCCAGGCTTCAAAGTACAGAGCGAGAATCTCTCTGCAGGGAATAAAGAAAATATTAA	2460
Qy	2461	CCCAATGAAACATCACTAGCTTCTCAAAGCTGAAAAAAGGTATATCACCAGTTGTT	2520
Db	2461	CCCAATGAAACATCACTAGCTTCTCAAAGCTGAAAAAAGGTATATCACCAGTTGTT	2520
Qy	2521	TCTGAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGTTTACAG	2580
Db	2521	TCTGAACATAGAAAAACAGATTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGTTTACAG	2580
Qy	2581	CCAAGTTCTACTTCTGAACTCTAGTGAATCTATGGAATCAACTACTAAACAAAAATAGACAGGAGAAAA	2640
Db	2581	CCAAGTTCTACTTCTGAACTCTAGTGAATCTATGGAATCAACTACTAAACAAAAATAGACAGGAGAAAA	2640
Qy	2641	TCAAGAGATTGATCAAAGCAAAAATGAAACCAAGTCTAAGGATTTCTTTTCAATTGAAAAAT	2700
Db	2641	TCAAGAGATTGATCAAAGCAAAAATGAAACCAAGTCTAAGGATTTCTTTTCAATTGAAAAAT	2700
Qy	2701	AGCAGCAGCAACTGTATACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA	2760
Db	2701	AGCAGCAGCAACTGTATACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA	2760
Qy	2761	ATACTTAGTAAACGGHAGCAAGAGGGHACTTGAAGTCACTTCCCAAGGGGTTTCAGACT	2820
Db	2761	ATACTTAGTAAACGGHAGCAAGAGGGHACTTGAAGTCACTTCCCAAGGGGTTTCAGACT	2820
Qy	2821	TCCAGCCCGCATGTAAACAAGAGAAGACGATTAAGAGAGAGAAGAAAGACGCAGCTGAG	2880
Db	2821	TCCAGCCCGCATGTAAACAAGAGAAGACGATTAAGAGAGAGAAGAAAGACGCAGCTGAG	2880
Qy	2881	CAAGTTAGGAAATCAACATTTGAATTCCTCAATGCAAGAGAGTTCAACCCAGTTCCTTCTCT	2940
Db	2881	CAAGTTAGGAAATCAACATTTGAATTCCTCAATGCAAGAGAGTTCAACCCAGTTCCTTCTCT	2940
Qy	2941	CAGCCAAAGGCTTCTACTAAGCCAACTTCACTCTCGGCCTCAAGACAACTAGCCCATCT	3000
Db	2941	CAGCCAAAGGCTTCTACTAAGCCAACTTCACTCTCGGCCTCAAGACAACTAGCCCATCT	3000
Qy	3001	ATGTTGGTGCATCAACAGCCAACTCCAGTTTATACTCAGCTGTTGTTGTTTGACCAAAAT	3060
Db	3001	ATGTTGGTGCATCAACAGCCAACTCCAGTTTATACTCAGCTGTTGTTGTTTGACCAAAAT	3060
Qy	3061	ATGATGTATCCAGTCCAGTGGAGCCAGGGGTGCAACTTTTATACCCAAATACCTATGACG	3120
Db	3061	ATGATGTATCCAGTCCAGTGGAGCCAGGGGTGCAACTTTTATACCCAAATACCTATGACG	3120
Qy	3121	CCCATGCCAGTGAATCAAGGCCAAGACATATAGACAGTACAAATATGCCCCAACAGCGG	3180
Db	3121	CCCATGCCAGTGAATCAAGGCCAAGACATATAGACAGTACAAATATGCCCCAACAGCGG	3180
Qy	3181	CAAGCCAGCATCATCAGATGCCATGATGCACCCAGCGTTCAGCAGCGGCCACCGATT	3240
Db	3181	CAAGCCAGCATCATCAGATGCCATGATGCACCCAGCGTTCAGCAGCGGCCACCGATT	3240



QY 3241 GCAGCCACCCACAGCTTACTCCAGGCAATATGTTGCCCTACAGTCCTCAGCAGTTCCTCA 3300  
DB 3241 GCAGCCACCCACAGCTTACTCCAGGCAATATGTTGCCCTACAGTCCTCAGCAGTTCCTCA 3300  
QY 3301 AATCAGCCCTTGTTCAGCATGTGCGACATATATCAGTCTCAGCATCCTCATGCTATAGT 3360  
DB 3301 AATCAGCCCTTGTTCAGCATGTGCGACATATATCAGTCTCAGCATCCTCATGCTATAGT 3360  
QY 3361 CTTGTAATACAGGGTAATGCTAGATGATGGACCAACCAACACAGCCCGAGCCTGGTTTA 3420  
DB 3361 CTTGTAATACAGGGTAATGCTAGATGATGGACCAACCAACACAGCCCGAGCCTGGTTTA 3420  
QY 3421 GTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACGCATGCGATGATCATGT 3480  
DB 3421 GTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACGCATGCGATGATCATGT 3480  
QY 3481 CCCAAATTACCATACAAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540  
DB 3481 CCCAAATTACCATACAAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540  
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCCCTGCAACCCACATCTCCACACCT 3600  
DB 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCCCTGCAACCCACATCTCCACACCT 3600  
QY 3601 CAGCCTTCAGTACCCCACTGGACAGCAGCAAGCCCAACATGTTGGAAGTCACTCTGCA 3660  
DB 3601 CAGCCTTCAGTACCCCACTGGACAGCAGCAAGCCCAACATGTTGGAAGTCACTCTGCA 3660  
QY 3661 CCCAGTCCTGTTACAGCACCATCAGACAGGCGCGCCAGGCTCCATCTGCGCCAGTCCA 3720  
DB 3661 CCCAGTCCTGTTACAGCACCATCAGACAGGCGCGCCAGGCTCCATCTGCGCCAGTCCA 3720  
QY 3721 CAGCAGCAGTACGCCATTTACCAACGCGGGCTTGGCCAACTCCACCTCCATGACACT 3780  
DB 3721 CAGCAGCAGTACGCCATTTACCAACGCGGGCTTGGCCAACTCCACCTCCATGACACT 3780  
QY 3781 GCCTCAACACGCGAGTGGCAGAGATGTTTCCAGAGCAACAAGACTGCTTTTACG 3840  
DB 3781 GCCTCAACACGCGAGTGGCAGAGATGTTTCCAGAGCAACAAGACTGCTTTTACG 3840  
QY 3841 ATCCATCTTCTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACAGTACT 3900  
DB 3841 ATCCATCTTCTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACAGTACT 3900  
QY 3901 CAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCCAACTGCGCCATGATGATG 3960  
DB 3901 CAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCCAACTGCGCCATGATGATG 3960  
QY 3961 CTAATGACGACACAGCAGCAGCGGTCCCAAGCGCGCTCGCTCAAAGTGCACTACAG 4020  
DB 3961 CTAATGACGACACAGCAGCAGCGGTCCCAAGCGCGCTCGCTCAAAGTGCACTACAG 4020  
QY 4021 CCCATTCCAGTCTCAGCAACAGCGCATTTCCCTATATAGCAGCAGCCTTCAGTACAGCC 4080  
DB 4021 CCCATTCCAGTCTCAGCAACAGCGCATTTCCCTATATAGCAGCAGCCTTCAGTACAGCC 4080  
QY 4081 CACCAACCAACAGCAGTGTGAGGCTGCGCTGAGGAAACCGAAAGGCCAAATTTCCCTCCTC 4140  
DB 4081 CACCAACCAACAGCAGTGTGAGGCTGCGCTGAGGAAACCGAAAGGCCAAATTTCCCTCCTC 4140  
QY 4141 CTTTCTACTGCTTCAACCACTGGAAGCAGACAGAAATCTAGAAATTTCAATTTTGTGTTT 4200  
DB 4141 CTTTCTACTGCTTCAACCACTGGAAGCAGACAGAAATCTAGAAATTTCAATTTTGTGTTT 4200  
QY 4201 TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAACAGTTCACTGCGAG 4260  
DB 4201 TAAATATATATGTTGATTTCTTGTAAACATCAATAGGAATGCTAACAGTTCACTGCGAG 4260  
QY 4261 TGGAGAGATCTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGGTATTTCCATAATTTCCA 4320  
DB 4261 TGGAGAGATCTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGGTATTTCCATAATTTCCA 4320

QY 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGGAACCTGGAAGTTATTT 4380  
DB 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGGAACCTGGAAGTTATTT 4380  
QY 4381 ATTTTTTAATACCCCTTGAAAGTCAATGAACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440  
DB 4381 ATTTTTTAATACCCCTTGAAAGTCAATGAACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440  
QY 4441 GATTCTTGCTCTATTAATGCTGCTAAAAAATAAAAAAAAAAAAA 4481  
DB 4441 GATTCTTGCTCTATTAATGCTGCTAAAAAATAAAAAAAAAAAAA 4481

RESULT 4  
US-10-802-228-1  
; Sequence 1, Application US/10802228  
; Publication No. US20050209178A1  
; GENERAL INFORMATION:  
; APPLICANT: Pulist, Stefan M  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF OBESITY  
; FILE REFERENCE: 825466-100151  
; CURRENT APPLICATION NUMBER: US/10/802,228  
; CURRENT FILING DATE: 2004-03-16  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 4481  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-802-228-1

Query Match 100.0%; Score 4481; DB 9; Length 4481;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCCGCTCTCACGTGTCCCTCCGCGGCCCGGG 60  
DB 1 ACCCCGAGAAAGCAACCCAGCGCGCGCCGCTCTCACGTGTCCCTCCGCGGCCCGGG 60  
QY 61 GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120  
DB 61 GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120  
QY 121 CACTTCGCTCCCAACCCGCGCCTCGGCGCGCCGCCCTCCGATGCGTCAAGCGGCCGCA 180  
DB 121 CACTTCGCTCCCAACCCGCGCCTCGGCGCGCCGCCCTCCGATGCGTCAAGCGGCCGCA 180  
QY 181 GCTCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGGAGCAGGTTGGGCC 240  
DB 181 GCTCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGGAGCAGGTTGGGCC 240  
QY 241 GGGTGGGCTCGCTCCAGCGCGCGCGGAGCGGGCGGGCGGGCGGGCGGGCGGCC 300  
DB 241 GGGTGGGCTCGCTCCAGCGCGCGCGGAGCGGGCGGGCGGGCGGGCGGGCGGCC 300  
QY 301 CCGGACCGTATCCCTCCGCGCCCTCCCGCGCGCCCGCCCGCCCGCCCTCCCTCCCGG 360  
DB 301 CCGGACCGTATCCCTCCGCGCCCTCCCGCGCGCCCGCCCGCCCGCCCTCCCTCCCGG 360  
QY 361 CAGAGTCCCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCAACCGGGCGGCCG 420  
DB 361 CAGAGTCCCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCAACCGGGCGGCCG 420  
QY 421 TTTCCGCGCGGCTCCCGCGGCTCTTGGTCTCGGGCGGCTCCCGCCCTTCGTCTGTC 480  
DB 421 TTTCCGCGCGGCTCCCGCGGCTCTTGGTCTCGGGCGGCTCCCGCCCTTCGTCTGTC 480  
QY 481 GTCCTTCTCCCTCCGCGAGCCCGGGCGCCCTCCCGCGCGCCAAACCCGCGCCTCCCGG 540  
DB 481 GTCCTTCTCCCTCCGCGAGCCCGGGCGCCCTCCCGCGCGCCAAACCCGCGCCTCCCGG 540  
QY 541 CTCGCGCGCGGTGCGTCCCGCGCGGTTCCGGGCTCTCTTGGGCGGCCCGGCTCCCGGC 600

Db 541 CTCGGGCCCCGTGCTCCCGCGCGTTCGGCGTCTCTTGGCGGCGCGCTCCCGC 600  
Qy 601 TGTCCCGCCCGCGGTGCGAGCGGTGTATGGGCCCTTCCACCATGTGCTGAAGCCCCAG 660  
Db 601 TGTCCCGCCCGCGGTGCGAGCGGTGTATGGGCCCTTCCACCATGTGCTGAAGCCCCAG 660  
Qy 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720  
Db 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720  
Qy 721 CAGCGCGCGCGCGCTCCCAATGTCCGAAGCCGCGGCGAGCGGCTTCTAGCTCG 780  
Db 721 CAGCGCGCGCGCGCTCCCAATGTCCGAAGCCGCGGCGAGCGGCTTCTAGCTCG 780  
Qy 781 CCGCGCGCGCGCTTCCGCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGT 840  
Db 781 CCGCGCGCGCGCTTCCGCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGT 840  
Qy 841 TCTCGTGGTGGCGGAGCTCCGCGCGGCGGAGCGCGCTGGGCGAGGTGGAAC 900  
Db 841 TCTCGTGGTGGCGGAGCTCCGCGCGGCGGAGCGCGCTGGGCGAGGTGGAAC 900  
Qy 901 AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAATATGAGG 960  
Db 901 AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAATATGAGG 960  
Qy 961 ATGGTTTCAATATCATCATCAGTTGTTGGTCCAAATGTGAAGTCAAGTGAATAATGGA 1020  
Db 961 ATGGTTTCAATATCATCATCAGTTGTTGGTCCAAATGTGAAGTCAAGTGAATAATGGA 1020  
Qy 1021 GGTATATATGAAGGAGTTTAAACTTACAGTCCGAGTGAATTTGTTGTTGTTGTTGTTG 1080  
Db 1021 GGTATATATGAAGGAGTTTAAACTTACAGTCCGAGTGAATTTGTTGTTGTTGTTGTTG 1080  
Qy 1081 GCACATGAAAGTACAGATCCAGTCCGGGCGGAAACGTGAAGAAATATGAGAGT 1140  
Db 1081 GCACATGAAAGTACAGATCCAGTCCGGGCGGAAACGTGAAGAAATATGAGAGT 1140  
Qy 1141 ATTTTCTTCAAAATGTTTCAAGTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200  
Db 1141 ATTTTCTTCAAAATGTTTCAAGTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200  
Qy 1201 GCAAAAGAGATGCTTTTACTGATCTGCTATCAGTCTAAAGTGAATGGCGAACAATA 1260  
Db 1201 GCAAAAGAGATGCTTTTACTGATCTGCTATCAGTCTAAAGTGAATGGCGAACAATA 1260  
Qy 1261 GAGAGGACTGAGCGCTGGATGAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1320  
Db 1261 GAGAGGACTGAGCGCTGGATGAGTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1320  
Qy 1321 TTGGAAATGACGTATCTAATGATGGGATCCCAATGATGTTTCGATATATGAAGAA 1380  
Db 1321 TTGGAAATGACGTATCTAATGATGGGATCCCAATGATGTTTCGATATATGAAGAA 1380  
Qy 1381 AATTATGGTGTAGTCTAGTATGATAGCAGTTTATCTTCTGATATACAGTGCCTTAGAA 1440  
Db 1381 AATTATGGTGTAGTCTAGTATGATAGCAGTTTATCTTCTGATATACAGTGCCTTAGAA 1440  
Qy 1441 AGAGATAACTCAGAGAAATTTTAAACCGGAGAGCAGGCAACAGTTAGCAGAGAA 1500  
Db 1441 AGAGATAACTCAGAGAAATTTTAAACCGGAGAGCAGGCAACAGTTAGCAGAGAA 1500  
Qy 1501 ATTGAGTCAAGTCCAGTACAAAGTCCAGTGGCCCTGGAAATGATGATAGAGTGAG 1560  
Db 1501 ATTGAGTCAAGTCCAGTACAAAGTCCAGTGGCCCTGGAAATGATGATAGAGTGAG 1560  
Qy 1561 GAAGAAATATACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCAACAGATAAAC 1620  
Db 1561 GAAGAAATATACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCAACAGATAAAC 1620  
Qy 1621 ACTAGGAAATATATATTTCTCTGGCAAGAAATAGAGAGTTCATATCTCTGGGA 1680  
Db 1621 ACTAGGAAATATATATTTCTCTGGCAAGAAATAGAGAGTTCATATCTCTGGGA 1680

Qy 1681 AGTGGGAGACAGATTTACCGGCTATGGCCAGCCTGGATCGGCTCCATGCCATCAAGA 1740  
Db 1681 AGTGGGAGACAGATTTACCGGCTATGGCCAGCCTGGATCGGCTCCATGCCATCAAGA 1740  
Qy 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATTTCTGGTTCAAGCAAAAGAGTAGTTAAT 1800  
Db 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATTTCTGGTTCAAGCAAAAGAGTAGTTAAT 1800  
Qy 1801 GGAGTGTTCCTGCGCCATCGCCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
Db 1801 GGAGTGTTCCTGCGCCATCGCCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
Qy 1861 CAGTCAGTCCCAACTCTCTTCACCTCGGCGAGCAACCCCTACAGCGCGCCCTCCAGG 1920  
Db 1861 CAGTCAGTCCCAACTCTCTTCACCTCGGCGAGCAACCCCTACAGCGCGCCCTCCAGG 1920  
Qy 1921 CCCCCCTCGGCGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGTCTCT 1980  
Db 1921 CCCCCCTCGGCGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGTCTCT 1980  
Qy 1981 GTCTCTACTATGCTTAAACGCGATGTTTCAAGAGGGCTTCCAAAGATGTCCCCAAAGGCC 2040  
Db 1981 GTCTCTACTATGCTTAAACGCGATGTTTCAAGAGGGCTTCCAAAGATGTCCCCAAAGGCC 2040  
Qy 2041 CAGCGACATCTTCGAAATCAGAGATTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100  
Db 2041 CAGCGACATCTTCGAAATCAGAGATTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100  
Qy 2101 GAATTTGTATCCCAACCCAGTGAAGCAGTACTCTCCAGTACAGAGGACAGT 2160  
Db 2101 GAATTTGTATCCCAACCCAGTGAAGCAGTACTCTCCAGTACAGAGGACAGT 2160  
Qy 2161 CCCTCGGGGGAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2220  
Db 2161 CCCTCGGGGGAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2220  
Qy 2221 CATAGACCCAGTCTCCAGACAGACAGTATTTGAAATACCCCGAGTGGGCCAGTTCTT 2280  
Db 2221 CATAGACCCAGTCTCCAGACAGACAGTATTTGAAATACCCCGAGTGGGCCAGTTCTT 2280  
Qy 2281 GCTTCTCCCAAGCTGTTATTTCAACTGAAGCTGTTGCCATGCTATTCCAGCTGCA 2340  
Db 2281 GCTTCTCCCAAGCTGTTATTTCAACTGAAGCTGTTGCCATGCTATTCCAGCTGCA 2340  
Qy 2341 TCTCTCTACCCCTGCTAGTCTCGATCGAACAGAGCTGTTACCCCTTCTAGTGAAGCTAAA 2400  
Db 2341 TCTCTCTACCCCTGCTAGTCTCGATCGAACAGAGCTGTTACCCCTTCTAGTGAAGCTAAA 2400  
Qy 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGGAATAAGAAATATATAA 2460  
Db 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGGAATAAGAAATATATAA 2460  
Qy 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAGTTGTT 2520  
Db 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAGTTGTT 2520  
Qy 2521 TCTGAACATAGAAAAACAGATTTGATGATTTAAAGAAATTTAAGAAATGATTTAGGTTACAG 2580  
Db 2521 TCTGAACATAGAAAAACAGATTTGATGATTTAAAGAAATTTAAGAAATGATTTAGGTTACAG 2580  
Qy 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAATAGAGAGGAGAAAAA 2640  
Db 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAATAGAGAGGAGAAAAA 2640  
Qy 2641 TCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTGTCTAAGGATTTCTTATTGAAAT 2700  
Db 2641 TCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTGTCTAAGGATTTCTTATTGAAAT 2700  
Qy 2701 AGCAGCAGCAACTGTACCTAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760  
Db 2701 AGCAGCAGCAACTGTACCTAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760

```
QY 2761 ATACTTAGTAAACACGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACT 2820
Db 2761 ATACTTAGTAAACACGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACT 2820
QY 2821 TCCAGCCAGCATGTAAACAAGAGAAAGACGATAGGAAGAGAAAGAGCGAGCTGAG 2880
Db 2821 TCCAGCCAGCATGTAAACAAGAGAAAGACGATAGGAAGAGAAAGAGCGAGCTGAG 2880
QY 2881 CAAGTTAGGAATCAACATTTGAATCCCAATGCAAAAGGAGTTCAACCCACGTTTCCTTCT 2940
Db 2881 CAAGTTAGGAATCAACATTTGAATCCCAATGCAAAAGGAGTTCAACCCACGTTTCCTTCT 2940
QY 2941 CAGCCAAAGCCCTTCTACTACCCCACTTCACTCGGCCCTCAAGCAACACCTAGCCCATCT 3000
Db 2941 CAGCCAAAGCCCTTCTACTACCCCACTTCACTCGGCCCTCAAGCAACACCTAGCCCATCT 3000
QY 3001 ATGGTGGGTCTCAACAGCCAACTCCAGTTTATATCTCAGCCCTGTTTGTGTCACCAAT 3060
Db 3001 ATGGTGGGTCTCAACAGCCAACTCCAGTTTATATCTCAGCCCTGTTTGTGTCACCAAT 3060
QY 3061 ATGATGTATCCAGTCCCAAGTCCAGCGGTCGCAACCTTTATACCCCAATACCTATGACG 3120
Db 3061 ATGATGTATCCAGTCCCAAGTCCAGCGGTCGCAACCTTTATACCCCAATACCTATGACG 3120
QY 3121 CCCATGCCAGTCAATCAAGCCNAGCATATAGAGCAGTACCAATATGCCCCCAACAGCGG 3180
Db 3121 CCCATGCCAGTCAATCAAGCCNAGCATATAGAGCAGTACCAATATGCCCCCAACAGCGG 3180
QY 3181 CAAGACCAAGCATCATCAGAGTGCATGATGACCCAGCGTCAAGAGCGGGCCCAACCGATT 3240
Db 3181 CAAGACCAAGCATCATCAGAGTGCATGATGACCCAGCGTCAAGAGCGGGCCCAACCGATT 3240
QY 3241 GCAGCCACCCCAACAGCTTACTCCACGCAATATGTTGCTTACAGTCCCTCAGCAGTTCCCA 3300
Db 3241 GCAGCCACCCCAACAGCTTACTCCACGCAATATGTTGCTTACAGTCCCTCAGCAGTTCCCA 3300
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCATGTCATAGT 3360
Db 3301 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCATGTCATAGT 3360
QY 3361 CCTGTAATACAGGGTAATGCTAGAAATGATGGCACCACCAACACAGCGCCCGCTGGTTTA 3420
Db 3361 CCTGTAATACAGGGTAATGCTAGAAATGATGGCACCACCAACACAGCGCCCGCTGGTTTA 3420
QY 3421 GTATCTTCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGAGCGCATGCGATGTCATGT 3480
Db 3421 GTATCTTCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGAGCGCATGCGATGTCATGT 3480
QY 3481 CCCAAATTCACATACAAAGAGAGCAAGCCCTTCTTCTACTTTGCGCATTTCCACGGGC 3540
Db 3481 CCCAAATTCACATACAAAGAGAGCAAGCCCTTCTTCTACTTTGCGCATTTCCACGGGC 3540
QY 3541 TCCCTTGTCTCAGCATATGCGCACCTTAACGCTACCTGACCCACCACTACTCCACACCT 3600
Db 3541 TCCCTTGTCTCAGCATATGCGCACCTTAACGCTACCTGACCCACCACTACTCCACACCT 3600
QY 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCAAATGTTGGAAGTCATCCTGCA 3660
Db 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCAAATGTTGGAAGTCATCCTGCA 3660
QY 3661 CCCAGTCTGTTCAGCACCATCAGCACCGCGGCTTCCATCTCGGCAGTCCA 3720
Db 3661 CCCAGTCTGTTCAGCACCATCAGCACCGCGGCTTCCATCTCGGCAGTCCA 3720
QY 3721 CAGCAGCAGTCAGCCATTTACACGCGGGGTTGGCCAACTCCACCTCCATGACACCT 3780
Db 3721 CAGCAGCAGTCAGCCATTTACACGCGGGGTTGGCCAACTCCACCTCCATGACACCT 3780
QY 3781 GCCTCCAAACGCGAGTCGCCACAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTACG 3840
Db 3781 GCCTCCAAACGCGAGTCGCCACAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTACG 3840
QY 3841 ATCCATCCTTCTCAGGTTTCAGCGCGGCGTATACCAACCCACCATGCGCCACGTACCT 3900
```

```
Db 3841 ATCCATCCTTCTCAGGTTTCAGCGCGGCGTATACCAACCCACCATGCGCCACGTACCT 3900
QY 3901 CAGGCTCATGTACAGTACAGGAATGTTCTCTTCTCATCCAACTGCGCCATGCGCCCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTACAGGAATGTTCTCTTCTCATCCAACTGCGCCATGCGCCCAATGATG 3960
QY 3961 CTAATGACGACACAGCCACCCGGCGTCCCAGCGCGCCCTCGCTCAAAGTGCACTACAG 4020
Db 3961 CTAATGACGACACAGCCACCCGGCGTCCCAGCGCGCCCTCGCTCAAAGTGCACTACAG 4020
QY 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAAGCC 4080
Db 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAAGCC 4080
QY 4081 CACCACCAACAGCAGTTGTAAGGCTGCCCTGAGAGAACCCGAAAGGCCAAATTCCTCTCTC 4140
Db 4081 CACCACCAACAGCAGTTGTAAGGCTGCCCTGAGAGAACCCGAAAGGCCAAATTCCTCTCTC 4140
QY 4141 CTTTCTACTGCTTCTACCAACTGGAAGCACAGAAAACCTAGAAATTTCAATTTATTTGTTTT 4200
Db 4141 CTTTCTACTGCTTCTACCAACTGGAAGCACAGAAAACCTAGAAATTTCAATTTATTTGTTTT 4200
QY 4201 TAAATATATATGTTGATTTCTTGTAACTCAATAGGAATGCTTAAAGTTCACTTGCGAG 4260
Db 4201 TAAATATATATGTTGATTTCTTGTAACTCAATAGGAATGCTTAAAGTTCACTTGCGAG 4260
QY 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAACTTGGGGCTATTCCATAAATCCA 4320
Db 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAACTTGGGGCTATTCCATAAATCCA 4320
QY 4321 TATGCTGTTTCAGAGTCCCGCAGTACCCAGTCTCTGTTGCCGAAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTCAGAGTCCCGCAGTACCCAGTCTCTGTTGCCGAAACTGGAAGTTATTT 4380
QY 4381 ATTTTAAATACCCCTTGAAGTTCATGAACACATCAGTAGCAAAAGAGTAACAAGAGT 4440
Db 4381 ATTTTAAATACCCCTTGAAGTTCATGAAGTTCATGAAGTTCATGAAGTTCATGAAGT 4440
QY 4441 GATTCTTGTCTCTATTACTGCTAAAAAATAAAAAAAAAAAAAA 4481
Db 4441 GATTCTTGTCTCTATTACTGCTAAAAAATAAAAAAAAAAAAAA 4481

RESULT 5
US-10-141-541-3
; Sequence 3, Application US/10141541
; Publication No. US20030167495A1
; GENERAL INFORMATION:
; APPLICANT: Pulat, Stefan M.
; TITLE OF INVENTION: SCA2 Knockout Animal and Methods of Use
; FILE REFERENCE: P-CE 5244
; CURRENT APPLICATION NUMBER: US/10/141,541
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/289,231
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4225
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27) ... (3884)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (222) ... (224)
; OTHER INFORMATION: ccs=Pro
US-10-141-541-3
```

Query Match 70.0%; Score 3138.2; DB 6; Length 4225;  
Best Local Similarity 86.2%; Pred. No. 0;

	Matches 3662; Conservative	1; Mismatches 473; Indels 112; Gaps 14;	
QY	222	CGCCGACGACAGGTGGCCCGGGGTGGCGCTCGCTCCAGCGCGCGCGCGCGAGCGGGCG	281
DB	71	CCCCGAGCCGCGCCCGCGCGGCTGGGCGCGCGCGCTCGCTCCGCGAGACGCGCG	130
QY	282	GGCGCGCGGTGGCGCGGCCCGGGGACGGTATCCCTCGCGCGCCCTCCCGCGCCCGGCC	341
DB	131	GGCGCGCGCGCGGGCGGGCGGTGGCGTATCCCTCCGCGCGCCCTCCCGCGCGCGGCC	190
QY	342	CGSCCCCTCCCTCCCGCGACAGCTCGCTCCCTCCGCTCAGACTGTTTTGGTAGCAA	401
DB	191	CGCGCCCTCCCGCGGGCGCGCTCGCACCGTGGCTCAGACTGTTTTGGTAGCAA	250
QY	402	CGGCAACGGCGGGCGCGCTTTTCGGCCCGGCTCCCGCGCGCTCCTTTGGTCTCGCGCGGCC	461
DB	251	CGGCCA-----CGGCGGTCCTCCCGCGCGCTCCCGCGCGCTCGGTGTCTCGCGGCC	304
QY	462	TCCCGCGCCCTTCGTGTGTCTCTTCCTCCCTTCGCCAGCGCCGGGGCGCTCCCGCGCG	521
DB	305	TCCCGCGCCCTTCGTGTGTCTTCGTGTCTCTG-----GCCCGCGCGGCCAC	352
QY	522	GCCAAACCGCGCTCCCGCTCGGGCGGTGGCGTCCCGCGCGCTTCGGGTCTCGGGCGCT	581
DB	353	GCGGCGCGCGCTTCGCCCGCGCGCGGTTCGCCGCTCCCGCGCGCTTCGGCGCTCTCTC	412
QY	582	---GGCGCGCCCGGCTCCCGCTGTCCCGCGCGGCTGCGAGCCGGTGTATGGGCGCCCT	638
DB	413	CTCGGCGGCGCCGGCACCGGCTGTCCCGCGCGGCGTGGAGCCGGTGTATGGGCGCT	472
QY	639	CACCATGTGCTGAAGCCCGACGACGACGACGACGACGACGACGACGACGACGACGACG	698
DB	473	CACCATGTGCTGAAGCC-----	490
QY	699	ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	758
DB	491	-----GCAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	532
QY	759	CGCGAGCGGCTTTAGCGTGGCGCGCGCGCGCTTCGCCGCTCTCGCTCTCGGTCTC	818
DB	533	CGCGGCGCTCTCTCGTGGCGCGGCGCGCGCGCGCTCGCGCGGCTGACCTCGGCTTC	592
QY	819	CTCGTCTCTCGGCCACCGGCTCCCTCTCGGTG-----GTGCGGCGACCTTCGGCGCGGGAG	875
DB	593	CGTGGTGC CGGCCCGCGCGCGCGGTGGGTCTTCTCGCGCGCGCGCGCGCGCGCGCG	652
QY	876	GCCGCGCTCGGCAGAGGTGAAACAGTAAACAAAGACTGCTCTCAGTCTACGATTTCTTT	935
DB	653	TCCCGGCTCGGCAGAGGTGCGAACAGTAGCAAGGACTGCTCTCAGCTACGATTTCTTT	712
QY	936	TGATGGAATCTATGCAATATGAGATGGTTTCAATATATATCTTACATCAGTCTGGCTCAA	995
DB	713	TGATGGAATCTATGCAACGTGAGATGGTTTCAATATATCTTACGTCAGTCTGGATCGAA	772
QY	996	ATGTGAAGTACAAGTGAATAATGAGGTATATATGAAGAGTTTTTAAATCTTACAGTCC	1055
DB	773	ATGTGAAGTACAAGTGAATAACGAGGCATATATGAAGGAGTTTTTAAACATACAGTCC	832
QY	1056	GAAGTGTGATTTGGTACTTTCATCGCGCACATGAGAAAGTACAGATCCAGTTCGGGGCC	1115
DB	833	TAAAGTGTGACTTGGTACTTGTATGCTGCACATGAGAAAGTACAGATCCAGTTCGGGGCC	892
QY	1116	GAAACGTGAAGAAATAATGAGAGTATTTTGGTTCAAATGTTTCAGACTTTGTTGGTGA	1175
DB	893	AAAACGTGAAGAAATAATGAGAGTGTTTTGGTTCAAATGTCAGACTTCGTTGGGTACA	952
QY	1176	GTTTAAAGATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAG	1235
DB	953	GTTTAAAGATACAGACTCCAGTTATGCACGGAGAGATGCTTTTACTCACTCTGCTCTCAG	1012
QY	1236	TGCTAAAGTGAATGGCGAAACAAAGAGAGGACCTGGAGCCCTGGGATGCGAGTGA	1295
DB	1013	CGAAAGGTGAATGGTGAGACACAGAGAGAGGACTCGGAGCCCTGGGATGCGAGGAGCT	1072

Qy	1296	CACAGCCAAATGAGGAACTTGAGGCTTTGGAAATAGCATGTCTAATTGGATGGATGCCAA	1355
Db	1073	CACGGCCAGCGAGGAGCTGGAG---CTGGAGAAATGATGTGTCTAATTGGATGGGACCCCAA	1129
Qy	1356	TGATATGTTTTCGATATAATGAAGAAATATGTGTAGTGTCTACGTATGATAGCAGTTT	1415
Db	1130	TGACATGTTTTCGATATAATGAAGAAATATGTGTGTGTGTCCACATATGATAGCAGTTT	1189
Qy	1416	ATCTTTCGTATACAGTGCCTTTAGAAAGAGATAACTCAGAAGAAATTTTAAACCGGAAGC	1475
Db	1190	ATCTTTCATATACGTTCTTTTGAAGAGGACAACCTCAGAAGAAATTTCTTAAACCGGAGGC	1249
Qy	1476	AAGGGCAAAACGAGTTAGAGAGAAATTTAGTCAAGTGCCAGTACAAAGCTCGAGTGCC	1535
Db	1250	AAGGGCAAAACGAGTTAGCAGAAGAAATTTGAATCCAGTGTCTCAGTACAAAGCTCGTGTCC	1309
Qy	1536	CCTTGGAAATGATAGTAGGAGTCAGGAAGAAATAACACAGCAGTTTCAGAGAAATTCACG	1595
Db	1310	CCTTGGAAATGATAGACCGGAGTCAGGAAGAAATAACACAGCAGTCAGAGAAATTCGACG	1369
Qy	1596	TGAACGTGAGGGGCACAGCAATAAACACTAGGGAAATAAATATATTTCTCTCGGACAAAG	1655
Db	1370	TGACCGGGAGGGCATGGCCCCAACACTAGGGACATAAATATATTTCTCTCGGACAAAG	1429
Qy	1656	AAATAGAGAAAGTCATATCTCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGCCAGCC	1715
Db	1430	AAACAGAGAAAGTCTATCTCTGGGGAAGTGGGAGACAGAGCTCAACCGGATGGCCAGCC	1489
Qy	1716	TGNAATCGGGCTCGATGCGCATCAAGATCCACTTCTCACATTTAGATTTCAACCCGAAATTC	1775
Db	1490	TGGGCCAGGGCTCCAATGCGGTCAAGAGCTGCTCTCTCACATTTCAAGATTTCAACCCGAAACGC	1549
Qy	1776	TGGTTACAGCAACAGAGTAGTTAATGAGAGTGTTCCTCTGGCCATCGCTTTGCCCATCTCC	1835
Db	1550	TGGTCTCAGACCAAGAGTAGTTAATGAGAGTGTTCCTCTGGCCATCGCTTTGCCCATCTCA	1609
Qy	1836	TTCTCTCTCGCCACCTTCTCGCTACCACTCAGGTCCCACTCTCTTCCACTCTCGGCAGC	1895
Db	1610	TTCTCTCTCGCCACCTTCTCGCTACCACTCAGGTCCCACTCTCTTCCACTCTCGGCAGC	1669
Qy	1896	CACCCCTACAGCGCGCCCTTCAGGCCCCCTCTCGGCCCATTCAGACCCCGGTCTCACCC	1955
Db	1670	CACCCATACAGCGCGCCCTTCAGGCCCCCTCTGAGGCCATTCAGACCCCGGTCTCACCC	1729
Qy	1956	CTCTGCTCATGGTTCTCCAGTCTGTCTCTACTATGCCTTAAGCGCATGCTTCAGAGG	2015
Db	1730	CTCTGCTCATGGTTCTCCAGTCTGTCTCTACTATGCCTTAAGCGCATGCTTCAGAGG	1789
Qy	2016	GCCTCCAGGATGTCCCAAGGCCACAGCACATCTCTCGAAATCACAGAGTTTCTGCTGG	2075
Db	1790	ACCCCAAGGATGTCTCAAAGGCACAGGCCACCTCTCGGAATCACAGAGTCTCTGCTGG	1849
Qy	2076	GAGGGGTTTCCATATCCAGTGCCCTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGC	2135
Db	1850	GAGAGGCTCCATGTCTAGTGCCTAGAAATTTGTATCCCAAACTCCCAAGTGAAGCAGC	1909
Qy	2136	TACTCTCCAGTAGCAAGGACAGTCCCTCGGGGGAAAGTGTCTCATCTAGTGTCTAGTGG	2195
Db	1910	TGCTCTCTCCAGTGGCAAGGACAGTCTCTGCAAGGGAAAGTGTGTCTCAGTGGTCTAGTGG	1969
Qy	2196	GGTTCCAGATTTATCCCTTAAACTCATAGACCAGGTCTCCACAGACAGAACAGTATTGG	2255
Db	1970	GGTTCCAGGTTTATCTCCAAAATCTCACAGACCAGGTCTCCAGGACAGCACATTTGG	2029
Qy	2256	AAATACCCCAAGTGGGCGAGTTCTTTGCTTTCTTCCCAAGCTGGTATTTATTTCAACTGAAGC	2315
Db	2030	AAACTCTCCCAAGGGGCTGTGCTTGTCTCTTCCCAAGCTGGCATCATCTCTGCAGAGC	2089
Qy	2316	TGTTGCCATGCCTATTCAGCTGCATCTCTTACGCCCTGTAGTCTCTGCATCGAACAGAGC	2375
Db	2090	GGTTTCCATGCTGTTTCCGCGCGCATCTCCGACTCTCTGCCAGCCCTGCATCCAAACAGAGC	2149

QY	2376	TGTTACCCCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCC	2433
DB	2150	ACTGACCCCATCTATTGAGGCAAAAGATTTCAGGCTTCAAGATCAGAGGCAGAACTCTCC	2209
QY	2436	TGCAGGGNAATAAGAAATATATAAACCAATGAACATCACCTAGCTTCTCAAAGCTGA	2495
DB	2210	TGCAGGGAGTAAGAGAAATGTTAAAGCAAGTGAACATCACCTAGCTTCTCAAAGCTGA	2269
QY	2496	AAACAAAGGTATATCACCAGTGTGTTCTGAAACATAGAAAAACAGATTGATGATTTAAAGAA	2555
DB	2270	CAACAAAGGTATGTCACAGTGTGTTCTGAACACAGAAAAACAGATTGATGATTTAAAGAA	2329
QY	2556	ATTTAAGAAATGATTTTAGTTTACAGCCAAAGTTCTACTCTGAAATCTATGGATCAACTACT	2615
DB	2330	GTTTTAAGAAATGATTTTAGTTTACAGCCAAAGCTCTACATCTGAATCTATGGATCAACTACT	2389
QY	2616	AAACAAAAATAGAGGGAGAAAAATCAGAGATTTGATCAAAGACAAAAATTGAAACCAAG	2675
DB	2390	AAGCAAAAAATAGAGAGGAGAAAAAGTCAAGAGATTTGATTAAGAGATATAACGGAAGCAAG	2449
QY	2676	TGCTAAGGATTCCTTTCAATTGAAAAAT-----AGCAGCAGCAACTGTACCAGTGGCAGCAG	2729
DB	2450	TGCTAAGGATAGTTTCATTGACAGCAGCAGCAGCAGCAGCACTGTGTACAGTGGCAGCAG	2509
QY	2730	CAAGCCGAATAGCCCCAGCAATTTCCCTTCAATATCTTAGTAACACGGAGCACAAGAGGGG	2789
DB	2510	CAAGACCAACAGCCTAGCATCTCCCTTCCATGCTTAGTAATGACAGCACAAGAGGGG	2569
QY	2790	ACCTGAGGTCACTTCCCAGGGGTTTCAGACTTCAGCCCGACGATGTAAAAACAGGAAAGA	2849
DB	2570	GCCTGAGGTCACTTCCCAGGGGTCGACACTTCAGCCCGACCTGCAAAACAGAGAAGGA	2629
QY	2850	CGATAAGGAAGAGAAGAACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAA	2909
DB	2630	TGACAGAGAGAGAAGAAAGACAAACAGACAGGTTAGGAAATCGACATTTGAATCCCAA	2689
QY	2910	TGCAAGGAGTTCAACCCACAGTTCCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTC	2969
DB	2690	TGCAAGGAGTTCAACCTCGTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAAGCTC	2749
QY	2970	ACCTCGGCCTCAAGCAAACTAGCGCAATCTATGTTGGTGGTTCATCAGCAGCGCAGTCCAGT	3029
DB	2750	ACCTCGGCCTCAAGCACAACCCAGCCCATCTATGTTGGTGGTTCATCAGCAGCGCAGTCCAGT	2809
QY	3030	TTATACTCAGCGCTGTTGTTTTTGACCAAAATATCATGTTCCAGTCCAGTCCAGTCCAGT	3089
DB	2810	GTACACTCAGCGCTGTGCTCTCGCACCCCAATATGATGTATCCCGTCCAGTGAAGCCCGGG	2869
QY	3090	CGTGCAACCTTTATACCCAATACCTATGAGCCCATGCCAGTGAATCAAGCCAAAGACATA	3149
DB	2870	CGTACAACCTTTATACCCAATACCTATGAGCCCATGCCAGTGAATCAAGCCAAAGACATA	2929
QY	3150	TAGAGC-----AGTACCAATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGC	3203
DB	2930	TAGAGCAGGTAAGATACCAATATGCCCAACAGCGGCAAGACCAATCATCAAAAGCAC	2989
QY	3204	CATGATGACCCAGCGTCAGCAGCGGGCCACCGATTGCGGCCACCCACCGAGTTACTC	3263
DB	2990	CATGATGACCCAGCGCTCCGGGAGGGCCACCCATCGTAGCCACCCCGCGCTTACTC	3049
QY	3264	CACGCAATATGTGCTACAGTCCCTCAGCAGTTCCCAAATCAGCGCCCTTGTTCAGCATGT	3323
DB	3050	CACTCAGTACGTTGCCTACAGCCCTCAGAGTTTCCCAATCAGCGCTTGTTCAGCATGT	3109
QY	3324	GCCACATTTACGTTCTACGCAATCCTCATGTCTATAGTCTGTAAATACAGGGTAATGCTAG	3383
DB	3110	GCCGCATTTACGTCTCAGCATCCTCATGTGTACAGTCTCTGTACAGTAAAGGTAATGCCAG	3169
QY	3384	AATGATGGCACCAACACAGCCCGCAGCTGGTTAGTATCTCTTTCAGCAACTCAGTA	3443
DB	3170	GATGATGGCACCAACAGCACATGTCAGCCGTGGTTTAGTGTCTTCTTCAGCTGCTCAGTT	3229
QY	3444	CGGGGCTCATGACGACAGCGCATGCGATGTATGCAATGTCCTCCAAATTCATATCAACCAAGGA	3503

[illegible]

RESULT 6  
US-09-925-298-15  
; Sequence 15, Application US/09925298  
; Publication No. US2002039764A1  
; GENERAL INFORMATION:





Db	1746	ATTCTTGCTGCTATTACTGCTRAAAAAAAAAAAAAAAAAAAAAA	1785
RESULT 7			
US-10-102-806-15			
; Sequence 15, Application US/10102806			
; Publication No. US20030054421A1			
; GENERAL INFORMATION:			
; APPLICANT: Rosen et al.			
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies			
; FILE REFERENCE: PA103PIC1			
; CURRENT APPLICATION NUMBER: US/10/102,806			
; CURRENT FILING DATE: 2002-03-22			
; PRIOR APPLICATION NUMBER: 09/925,298			
; PRIOR FILING DATE: 2001-08-10			
; PRIOR APPLICATION NUMBER: PCT/US00/05881			
; PRIOR FILING DATE: 2000-03-08			
; PRIOR APPLICATION NUMBER: 60/124,270			
; PRIOR FILING DATE: 1999-03-12			
; NUMBER OF SEQ ID NOS: 846			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 15			
; LENGTH: 2006			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (862)			
; OTHER INFORMATION: n equals a,t,g, or c			
; NAME/KEY: misc feature			
; LOCATION: (1006)			
; OTHER INFORMATION: n equals a,t,g, or c			
US-10-102-806-15			
Query Match			
Best Local Similarity 38.6%; Score 1730.6; DB 5; Length 2006;			
Matches 1768; Conservative 2; Mismatches 8; Indels 2; Gaps 3;			
Qy	2710	AACGTGTACAGTGGCAGCAGCAAGCGGAATAGCCCGCAGCATTTCCCCCTTCAATACTTAGT	2769
Db	6	AGCTGTACCAGTGGCAGCAGCAAGCGGAATAGCCCGCAGCATTTCCCCCTTCAATACTTAGT	65
Qy	2770	AACACGGAGCACAAGAGGGGACCTGAGTCACTTCCCAAGGGGTTTCAGACTTCCAGCCCA	2829
Db	66	AACACGGAGCACAAGAGGGGACCTGAGTCACTTCCCAAGGGGTTTCAGACTTCCAGCCCA	125
Qy	2830	GCATGTAACAGAGAAAGACGATAAGAGAGAGAAAGACGACGCTGAGCAAGTTAGG	2889
Db	126	GCATGTAACAGAGAAAGACGATAAGAGAGAGAAAGACGACGCTGAGCAAGTTAGG	185
Qy	2890	AAATCAACATTGAATCCCAATGCAAGAGGATTCAACCCACAGTTCTCTCTCAGCCAAAG	2949
Db	186	AAATCAACATTGAATCCCAATGCAAGAGGATTCAACCCACAGTTCTCTCTCAGCCAAAG	245
Qy	2950	CCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCCACTATGTTGGGT	3009
Db	246	CCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGCCCACTATGTTGGGT	305
Qy	3010	CATCAACAGCCCACTCGAGTTTATCTAGCCCTGTTTGTGACCAAAATATGATGAT	3069
Db	306	CATCAACAGCCCACTCGAGTTTATCTAGCCCTGTTTGTGACCAAAATATGATGAT	365
Qy	3070	CCAGTCCCAGTGAGCCCAAGGGTGCAACTTTATACCCCAATACCTATGACGCCCATGCCA	3129
Db	366	CCAGTCCCAGTGAGCCCAAGGGTGCAACTTTATACCCCAATACCTATGACGCCCATGCCA	425
Qy	3130	GTGAATCAAGCCCAAGACATATAGAGC-----AGTACCAATATGCCCCCAACAGCGGCAA	3183
Db	426	GTGAATCAAGCCCAAGACATATAGAGCGGTAAAGTACCAAAATATGCCCCCAACAGCGGCAA	485
Qy	3184	GACCAGCATCATCAGAGTGCCATGATGACCCAGGCTCAGCAGGGGCCCAACCGATTGCA	3243
Db	486	GACCAGCATCATCAGAGTGCCATGATGACCCAGGCTCAGCAGGGGCCCAACCGATTGCA	545
Qy	3244	GCCACCCACACAGCTTATCTCCACGCAATATGTTGCTACAGTCTCTCAGCAGTTCCCAAAAT	3303
Db	546	GCMACCCACACAGCTTATCTCCACGCAATATGTTGCTACAGTCTCTCAGCAGTTCCCAAAAT	605
Qy	3304	CAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCTCTCATGTCTATAGTCTCT	3363
Db	606	CAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCTCTCATGTCTATAGTCTCT	665
Qy	3364	GTAATACAGGGTAATGCTAGAAATCATGCGCACCAACACACAGCCCGAGCTGGTTTAGTA	3423
Db	666	GTAATACAGGGTAATGCTAGAAATCATGCGCACCAACACACAGCCCGAGCTGGTTTAGTA	725
Qy	3424	TCCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCGCATGGATGTATGCAATGTCTCC	3483
Db	726	TCCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCGCATGGATGTATGCAATGTCTCC	785
Qy	3484	AAATTACCATACAAAGGAGACAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	3543
Db	786	AAATTACCATACAAAGGAGACAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	845
Qy	3544	CTTGCTCAGCAGTATGCGC-ACCCTAAGCTACCTGCACCACCATACTCCACACCCCTCA	3602
Db	846	CTTGCTCAGCAGTATGNGCRACCTTAACGCTACCTGACCACCATACTCCACACCCCTCA	905
Qy	3603	GCCTTTCAGCTACCCCCACTGGACAGCAGCAAGCAACATGGTGGAAAGTCACTCTGCAACC	3662
Db	906	GCCTTTCAGCTACCCCCACTGGACAGCAGCAAGCAACATGGTGGAAAGTCACTCTGCAACC	965
Qy	3663	CAGTCTCTGTTTTCAGCAACCATCAGCAGCGCGCCAGGCT-CTCCATCTGGGCCAGTCCAC	3721
Db	966	CAGTCTCTGTTTTCAGCAACCATCAGCAGCGCGCCAGGCTNCTCATCTGGCCAGTCCAC	1025
Qy	3722	AGCAGCAGTTCAGCCATTTTACCACGCGGGGCTTGGCCAACTCCACCTCCATGACACCTG	3781
Db	1026	AGCAGCAGTTCAGCCATTTTACCACGCGGGGCTTGGCCAACTCCACCTCCATGACACCTG	1085
Qy	3782	CCTCCAAACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGAGTCTCTTTACGA	3841
Db	1086	CCTCCAAACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGAGTCTCTTTACGA	1145
Qy	3842	TCCATCTCTTTCAGCGCGGCTTACCAACCCACCCACATGGCCACGCTACCTC	3901
Db	1146	TCCATCTCTTTCAGCGCGGCTTACCAACCCACCCACATGGCCACGCTACCTC	1205
Qy	3902	AGGCTCATGTACAGTTCAGGAATGTTCTTCTCATCCAACTGCCCATGCGCCAAATGATGC	3961
Db	1206	AGGCTCATGTACAGTTCAGGAATGTTCTTCTCATCCAACTGCCCATGCGCCAAATGATGC	1265
Qy	3962	TAATGACGACACAGCCACCCGCGGTCCTCCAGCGCGGCTCCGCTCAAAGTGCACCTACAGC	4021
Db	1266	TAATGACGACACAGCCACCCGCGGTCCTCCAGCGCGGCTCCGCTCAAAGTGCACCTACAGC	1325
Qy	4022	CCATTCOAGTCTCGAACACAGCGCATTTCCCTATATAGCGACCCCTTCAGTACAAGCCC	4081
Db	1326	CCATTCOAGTCTCGAACACAGCGCATTTCCCTATATAGCGACCCCTTCAGTACAAGCCC	1385
Qy	4082	ACCACCAACAGCAGTTGTAGGCTGCCCTGGAGAAACCGAAGGCGCAAAATCCCTCTCTCC	4141
Db	1386	ACCACCAACAGCAGTTGTAGGCTGCCCTGGAGAAACCGAAGGCGCAAAATCCCTCTCTCC	1445
Qy	4142	CTTCTACTGCTTCTACCAACTGGAAGCAGACAGAAAATAGAAATTTCAATTTTGTGTTTTT	4201
Db	1446	CTTCTACTGCTTCTACCAACTGGAAGCAGACAGAAAATAGAAATTTCAATTTTGTGTTTTT	1505
Qy	4202	AAAATATATATGTTGATTTCTTGTAAACATCCAAATAGGAATGCTAACAGTTCACTTGCAGT	4261
Db	1506	AAAATATATATGTTGATTTCTTGTAAACATCCAAATAGGAATGCTAACAGTTCACTTGCAGT	1565
Qy	4262	GGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACTTGGGGGCTATTCCATAATTTCCAT	4321
Db	1566	GGAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACTTGGGGGCTATTCCATAATTTCCAT	1625







REGISTRATION NUMBER: 33,977  
REFERENCE/DOCKET NUMBER: 232.00010101  
TELEPHONE: 612/305-1220  
TELEFAX: 612/305-1228  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 516 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-750-323-1

Query Match 11.1%; Score 497.4; DB 8; Length 516;  
Best Local Similarity 99.8%; Pred. No. 9.4e-115;  
Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

392 TTGGTAGCAACGCAACGGCGGCGGCTTCGGCCCGGCTCCCGCGGCTCCTTGGTC 451  
1 TTGGTAGCAACGCAACGGCGGCGGCTTCGGCCCGGCTCCCGCGGCTCCTTGGTC 60  
452 TCGGCGGCGCTCCCGGCGGCTTCGTGTGTCTTCTCCCTTCGCGAGCCCGGCGGCC 511  
61 TCGGCGGCGCTCCCGGCGGCTTCGTGTGTCTTCTCCCTTCGCGAGCCCGGCGGCC 120  
512 CTCGCGCGGCGCAACCGCGGCTCCCGCTCGCGCGCGTGTCCCGCGCGTTCGG 571  
121 CTCGCGCGGCGCAACCGCGGCTCCCGCTCGCGCGCGTGTCCCGCGCGTTCGG 180  
572 GCGTCTCTTGGCGGCGCGGCTCCCGCTGTCCCGCGCGGCTGTCCCGCGGCTGTATG 631  
181 GCGTCTCTTGGCGGCGGCTCCCGCTGTCCCGCGCGGCTGTCCCGCGGCTGTATG 240  
632 GGCGCTTCACTGTGTCTGAAGCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 691  
241 GGCGCTTCACTGTGTCTGAAGCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 300  
692 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 751  
301 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360  
752 AGCCCGCGCGCGCGGCTTCTAGCGTGTCCCGCGCGGCTTCGCGCGTCTCGTCT 811  
361 AGCCCGCGCGCGGCTTCTAGCGTGTCCCGCGCGGCTTCGCGCGTCTCGTCT 420  
812 CGGTCTCTCTGTCTCGGCGCAGCGCTCCCTCTCGGTGTGTGTGTGTGTGTGTGTGT 871  
421 CGGTCTCTCTGTCTCGGCGCAGCGCTCCCTCTCGGTGTGTGTGTGTGTGTGTGTGT 480  
872 GGAGGCGCGCGCTGGGCGAG 890  
481 GGAGGCGCGCGCTGGGCGAG 499

RESULT 11  
US-10-956-157-3976/c  
Sequence 3976, Application US/10956157  
Publication No. US20050118625A1  
GENERAL INFORMATION:  
APPLICANT: Wyeth  
APPLICANT: Mounts, William  
TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
FILE REFERENCE: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES  
CURRENT APPLICATION NUMBER: US/10/956,157  
CURRENT FILING DATE: 2004-10-04  
NUMBER OF SEQ ID NOS: 319805  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3976  
LENGTH: 682  
TYPE: DNA

```

; ORGANISM: Homo sapiens
US-10-956-157-3976

Query Match      10.1%; Score 452; DB 9; Length 682;
Best Local Similarity 90.5%; Pred. No. 3.3e-103;
Matches 516; Conservative 0; Mismatches 0; Indels 54; Caps 1;

QY 3335 AGTCTCAGCATCCTCATGCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATCATCGCAC 3394
DB 681 AGTCTCAGCATCCTCATGCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATCATCGCAC 622
QY 3395 CACCAACACACGCCCGCAGCCCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATG 3454
DB 621 CACCAACACACGCCCGCAGCCCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATG 562
QY 3455 AGCAGAGCGCATGCGATGTATGCAATGTGCCAAATTCACATACAAAGGAGACAAAGCCCTT 3514
DB 561 AGCAGAGCGCATGCGATGTATG-----TTTTCACGGGCTCCCTTCTCAGCAGTATGCGCACCCCTAACCGCTA 3574
QY 3515 CTTCTTACTTTTGCCATTTTCCACGGGCTCCCTTCTCAGCAGTATGCGCACCCCTAACCGCTA 3574
DB 540 -----TTTTCACGGGCTCCCTTCTCAGCAGTATGCGCACCCCTAACCGCTA 496
QY 3575 CCCTGCAACCCACATACCTCCACACCCCTCAGCGCTTCAGCTACCCCGACCTGGACAGCAGCAAA 3634
DB 495 CCCTGCAACCCACATACCTCCACACCCCTCAGCGCTTCAGCTACCCCGACCTGGACAGCAGCAAA 436
QY 3635 GCCAACAATGTGGAAGTCATCTCTGCAACCCAGTCTCTGTTTCAGCAGCAGTATCAGCACAGGCGG 3694
DB 435 GCCAACAATGTGGAAGTCATCTCTGCAACCCAGTCTCTGTTTCAGCAGCAGTATCAGCACAGGCGG 376
QY 3695 CCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCCATTTACACCGCGGGGCTTG 3754
DB 375 CCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCCATTTACACCGCGGGGCTTG 316
QY 3755 CGCCAACTCCACCCCTCCATGACACCTCGCTCCCAACAGCGCAGTCCGCCACAGAATAGTTTCC 3814
DB 315 CGCCAACTCCACCCCTCCATGACACCTCGCTCCCAACAGCGCAGTCCGCCACAGAATAGTTTCC 256
QY 3815 CAGCAGCACACAGACTGTCTTTACGATCCATCTCTTCAGGTTTCAGCGCGGCGTATACCA 3874
DB 255 CAGCAGCACACAGACTGTCTTTACGATCCATCTCTTCAGGTTTCAGCGCGGCGTATACCA 196
QY 3875 ACCCACCACCATGGCCCGCAGTACCTCAGG 3904
DB 195 ACCCACCACCATGGCCCGCAGTACCTCAGG 166

RESULT 12
US-10-956-157-9211
; Sequence 9211, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES.
; FILE REFERENCE: 031896-043000 (AW 101081)
; CURRENT APPLICATION NUMBER: US/10/956.157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9211
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-9211

```

	Qy	812	CGGTCCTCCTCGTCCGCCACGGCTCCTCTCGGTGTCGCAGCACTTCGCGCGGG	87
	Dd	421	CGGTCTCCTCTCGTCCGCCACGGCTCCTCTCGGTGTCGCAGCGACTTCGCGCGGG	480
	Qy	872	GGAGGCCCGGCTGGGCAG	890
	Dd	481	GGAGGCCCGGCTGGGCAG	499

RESULT 11  
US-10-956-157-3976/c  
; Sequence 3976, Application US/10956157  
; Publication No. US20050118625A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOC  
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES  
; FILE REFERENCE: 031896-043000 (AM 101081)  
; CURRENT APPLICATION NUMBER: US/10/956,157  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 319805.  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3976  
; LENGTH: 682  
; TYPE: DNA



```
QY 181 GTCCTCTGAGATCCCGCGTGGCCACCGAGTCTCGCCGCTTCGCGCAGCCAGGTGGCCC 240
  |||||
Db 292 ACTCTCGAAATCCCGCGATTAACCAACCGAATCTGGCG- TTGCGCGCAACCAATAACCC 350
  |||||
QY 241 GSGTGGCGTCTCGTCCAGCGCGCGGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCG 300
  |||||
Db 351 GAATAACGCTCGCTCCAAACGACGCGACGAAACGAAACGAAACGACGATTAACGCGACC 410
  |||||
QY 301 CGGGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGG 360
  |||||
Db 411 CGGAACCGTATCCCTCCGCGCGCCCTCTCTCGCGCGACCGCGACCGCGCGCGCGCG 470
  |||||
QY 361 CAGAGCTCGCCTCCCTCCGCGCTCAGACTGTTTGTAGCAACGCGCAACGCGCGCGCGG 420
  |||||
Db 471 CAAACTCGCCTCCCTCCGCGCGCTCAAACTATTTTAATAACAACGACAAACGACGACGCG 530
  |||||
QY 421 TTTGCGCGCGGCTCCGCGGCGCTCTGCTCGGCGGCGCTCCCGCGCGCGCTTCGTCGTC 480
  |||||
Db 531 TTTGACCGCGACTCCCGAGACTCTTAATCTCGACGAACTCCCGCGCGCGCTTCGTCGTC 590
  |||||
QY 481 GTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
  |||||
Db 591 CTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 650
  |||||
QY 541 CTCGCGCGCGGTCGCTCCCGCGCGCGCTTCGCGCGCTCTCTTGGCGCGCGCGCGCT 594
  |||||
Db 651 CTCGACGCGCGCGCTCCCGCGCGCTTCGAGCTCTCTTAACGCGCGCGACT 704
  |||||
```

## RESULT 15

```
US-10-363-483A-23185/c
; Sequence 23185, Application US/10363483A
; Publication No. US2005006401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 23185
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 23185
US-10-363-483A-23185
```

```
Query Match          9.8%; Score 441.2; DB 9; Length 704;
Best Local Similarity 85.0%; Pred. No. 1.8e-100;
Matches 505; Conservative 0; Mismatches 88; Indels 1; Gaps 1;

QY 1 ACCCGCGAGAAAGCAACCCAGCGCGCGCGCGCTCTCAGCTGCTCCCTCCCGCGCGCGG 60
Db 593 ACCCGGAAAAAACAACCCAGCGCGCGCGCTCTCAGTATCCCTCCCGACCCCGAA 534
  |||||
QY 61 GCACCTCTCAGTCTGCTCTCCGCTCGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG 120
Db 533 ACCACCTCAGTTTCTACTTCCGCTTAACCCCTCCGACTTCGGATAAAAATCCCTATCCG 474
  |||||
QY 121 CACTTCGCTTCCACACCGCGCGCTTCGCGCGCGCGCGCTCCGATGCGCTCAGCGCGCGCA 180
Db 473 CACTTCGCTTCCACACCGCGCGCTTCGACGCGCGCGCGCTCCGATAGCTCAACGACGCA 414
  |||||
QY 181 GTCCTCTGGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCGCGCGCGCGCG 240
Db 413 ACTCCTCGAAATCCCGCGGATAACCGGAATCTCGCGG-TTCGCGCGCAACCAATAACCC 355
  |||||
```

```
QY 241 GSGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGCGGCGGCGGCTGGCGCGGCC 300
  |||||
Db 354 GAATAACGCTCGCTCCAAACGACGCGACGAAACGAAACGAAACGACGATTAACGCGACC 295
  |||||
QY 301 CGGGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGG 360
  |||||
Db 294 CGGAACCGTATCCCTCCGCGCGCCCTCTCTCGCGCGACCGCGACCGCGCGCTCCCTCCCGA 235
  |||||
QY 361 CAGAGCTCGCCTCCCTCCGCGCTCAGACTGTTTGTAGCAACGCGCAACGCGCGCGCGG 420
  |||||
Db 234 CAAACTCGCCTCCCTCCGCGCGCTCAAACTATTTTAATAACAACGACAAACGACGACGCG 175
  |||||
QY 421 TTTGCGCGCGGCTCCGCGGCGCTCTGCTCGGCGGCGCTCCCGCGCGCGCTTCGTCGTC 480
  |||||
Db 174 TTTGACCGCGACTCCCGAGACTCTTAATCTCGACGAACTCCCGCGCGCGCTTCGTCGTC 115
  |||||
QY 481 GTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
  |||||
Db 114 CTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 55
  |||||
QY 541 CTCGCGCGCGGTCGCTCCCGCGCGCGCTTCGCGCGCTCTCTTGGCGCGCGCGCGCT 594
  |||||
Db 54 CTCGACGCGCGCGCTCCCGCGCGCTTCGAGCTCTCTTAACGCGCGCGACT 1
  |||||
```

Search completed: December 23, 2005, 06:06:39  
Job time : 3078 secs

**This Page Blank (uspto)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	81.4	1.8	13672	7	US-11-055-035-25	Sequence 2, Appli
2	80.6	1.8	255	6	US-10-909-122-1963	Sequence 1963, Ap
C 3	70.2	1.6	5679	7	US-11-075-185-36	Sequence 36, Appl
4	70.2	1.6	78869	7	US-11-075-185-1	Sequence 1, Appli
5	69.5	1.6	1867	7	US-11-043-590-45	Sequence 45, Appl
6	69.2	1.5	1310	7	US-11-137-671-13	Sequence 13, Appl
7	69.2	1.5	3263	7	US-11-137-671-15	Sequence 15, Appl
8	69.2	1.5	4286	7	US-11-137-671-14	Sequence 14, Appl
C 9	69.2	1.5	11070	7	US-11-075-185-34	Sequence 34, Appl
10	68.4	1.5	3507	6	US-10-509-422-1	Sequence 1, Appli
11	68.4	1.5	3704	6	US-10-509-422-3	Sequence 3, Appli
12	68.4	1.5	20945	6	US-10-995-561-13463	Sequence 13463, A
13	68.4	1.5	23082	6	US-10-995-561-13457	Sequence 13457, A
14	68.4	1.5	94510	6	US-10-995-561-13332	Sequence 13332, A
C 15	68.4	1.5	14248	6	US-10-995-561-13381	Sequence 13381, A
16	67.4	1.5	96128	6	US-10-995-561-13197	Sequence 13197, A
C 17	66.8	1.5	2276	6	US-10-131-826A-9	Sequence 9, Appli
C 18	66.4	1.5	16082	6	US-10-995-561-13485	Sequence 13485, A
C 19	66.4	1.5	23894	6	US-10-995-561-13320	Sequence 13320, A
C 20	66.4	1.5	46215	6	US-10-995-561-13483	Sequence 13483, A
C 21	66.4	1.5	268685	6	US-10-933-025-22	Sequence 22, Appl
C 22	65.8	1.5	171936	6	US-10-933-025-24	Sequence 24, Appl
C 23	65.6	1.5	1433	7	US-11-090-351-1	Sequence 1, Appli



```
RESULT 2
US-10-909-125-1963
; Sequence 1963, Application US/10909125
; Publication No. US200502612181
; GENERAL INFORMATION:
; APPLICANT: Esau, Christine
; APPLICANT: Lollo, Bridget
; APPLICANT: Bennett, C. Frank
; APPLICANT: Freier, Susan M.
; APPLICANT: Griffey, Richard H.
; APPLICANT: Baker, Brenda F.
; APPLICANT: Vickers, Timothy
; APPLICANT: Marcusson, Eric G.
; APPLICANT: Koller, Erich
; APPLICANT: Swayze, Eric
; APPLICANT: Jain, Ravi
; APPLICANT: Bhat, Balkrishen
; APPLICANT: Peralta, Eigen
; TITLE OF INVENTION: Oligomeric Compounds And Compositions For Use In Modulation
; FILE REFERENCE: ISIS0080-100 (CORE0016US)
; CURRENT APPLICATION NUMBER: US/10/909,125
; CURRENT FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: US 60/492,056
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/516,303
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: US 60/531,596
; PRIOR FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/562,417
; PRIOR FILING DATE: 2004-04-14
; NUMBER OF SEQ ID NOS: 2184
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1963
; LENGTH: 255
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-909-125-1963

Query Match      1.8%; Score 80.6; DB 6; Length 255;
Best Local Similarity 65.0%; Pred. No. 3.7e-08;
Matches 119; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 585 GCGCCGCGCTCCCGCTGTCCCGCGCGGTGCGAGCGGTGTATGGGCCCTCACCAT 644
Db 32 GCGGGGATCAGCGCTGCGCGGCGGTTCGCTTCAGACACCCCATCAACGCCAGCACG 91

Qy 645 GTCGTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 704
Db 92 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 151

Qy 705 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 764
Db 152 GTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 211

Qy 765 CGG 767
Db 212 TGG 214

RESULT 3
US-11-075-185-36/c
; Sequence 36, Application US/11075185
; Publication No. US200502664341
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099_03
; CURRENT APPLICATION NUMBER: US/11/075,185
```

```
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-36

Query Match      1.6%; Score 70.2; DB 7; Length 5679;
Best Local Similarity 45.6%; Pred. No. 2.6e-05;
Matches 397; Conservative 0; Mismatches 468; Indels 6; Gaps 4;

Qy 23 GCGCGCGCGCTCCTCAGGTGTCCTCCCGCCCCGGGGCCACCTCAGTTCTGTGTTCCG 82
Db 4121 GCGCGCGTGGCCCGCAGCAGCGCCCGCGCCCGCCAGCGCGTCCACGCCCTTCG 4062

Qy 83 TCTCACCCCTCCGACTTCCGGTAAGAGTCCCTATCCGCACCTCCGCTCCACCC-CCGC 140
Db 4061 TCCGCGCGGTGCGCACCGCGCGCGGTCCACACAGAGCTCCGTGCGCGCCAGTGC 4002

Qy 141 GCCTCGGCGCGCGCGCTCCGATGCGCTCAGCGCGCGCGAGCTCTCGAGT--CCCCG 198
Db 4001 GCGCGCGGAGCCACCTCTGCACAGGCGCAGCGCGCGCGCACCGCGCGTGCACGCC 3942

Qy 199 GTGCGCACCGAGTCTCGCGCTTCGCGCAGCAGGTGCGCGGTGCGGTGCGTCCAG 258
Db 3941 GCGACGACCGCGCTGCATCCCGCGCTCAGGTCCACAGCAGCGCGCGCGCGCGCC 3882

Qy 259 CGGCGCGCGCGCGGAGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 318
Db 3881 GCGCGGTCCGCGAGCGCTCCAGCAGTCTCCGCGAGCGCGCGCGCGCGCGCGCGCG 3822

Qy 319 GCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 378
Db 3821 GCGCGCGGAGCGCACCGTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3762

Qy 379 GCCTCAGACTGTTTGTAGCAACGCGAAACGCGCGCGCGCGCGGTTCGCGCGCGCG 438
Db 3761 GCCTCCAGCGCACGTCCTCCACGCGACCGCGTACATCGGCTCCGCGCGCGCGCG 3702

Qy 439 CGGCTCTGTCTCGCGCGCGCTCCCGCGCGCTTCGTGCGTGTCTTCTCCCCCTGCC 498
Db 3701 GACCGGTCTGCGCGCGGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3642

Qy 499 AGCCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 558
Db 3641 TGCGCGTCCGCGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3582

Qy 559 CCGCGCGTTCGCGGTCTCTTGGCGCGCGCGGTTCGCGGTGTCCCGCGCGCGCGGTG 618
Db 3581 GCGATGCGCGCGCGCGTCTGCTCGCGCGCGGTTCGCGGTACAGC-GCCACGTCGTCACGC 3523

Qy 619 GAGCGGTGTATGGCCCTCCACCATGTGCGTGAAGCCCGCGCGCGCGCGCGCGCG 678
Db 3522 GAACGGCACCGACCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3463

Qy 679 CAGCAACAGCAGCAGCAACAGCAGCAGCAG- CAGCAGCAGCAGCGCGCGCGCGCG 737
Db 3462 CTGCAGCGCGCATCGAAGAGCGCGGATGGAGCCCGAAGCGCTGCTCGGTGCGCGCG 3403

Qy 738 TGCCAATGTCCGAAGCCCGCGCGCGCGCGCTTCTAGCGTTCGCGCGCGCGCGCGCTTC 797
Db 3402 AGCGTCCGCGCGCGCGTACAGCGGTACAGCGTACAGGTTCGCGCGCGCGCGCGCG 3343

Qy 798 GCGGTCTCTGTCCTGCTCTCTCGTCTCTCGGCGCGCGCTCCCTCTCGGTGCGCGCG 857
Db 3342 CCCCTGGAACGCGCGCGCGGTGGCGGTAGCGGTGCTCGGACAGCTGCTCGTACCGCGCG 3283
```

```
QY 858 GACCTCCGGCGGGAGGCGCGGCTGGGC 888
D  |||||
D 3282 GAGCGCCACCGCTCCGGCGCGGGCGGC 3252
    |||||

RESULT 4
US-11-075-185-1/c
; Sequence 1, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 78869
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-1

Query Match 1.6%; Score 70.2; DB 7; Length 78869;
Best Local Similarity 45.6%; Pred. No. 8.8e-05;
Matches 397; Conservative 0; Mismatches 468; Indels 6; Gaps 4;

QY 23 GCGCGCGCGCTCCTCAGCTGTCCTCCGCGCGCGCGCGCGCGCGCTTCTGCTTCG 82
D  |||||
D 30352 GCGCGGTGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30293

QY 83 TCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCGACCTCCGCTCCACCC--CGGC 140
D  |||||
D 30292 TCCGCGCGGTGCGCACCGCGAGCGGTACACACAGAGTCTCGTCGCGCGCGAGTGC 30233

QY 141 GCTCGCGCGCGCGCGCTCCGATGCGTACAGGCGCGCGAGTCTCTCGGAGT--CGCGG 198
D  |||||
D 30232 GCGCGCGCGAGCACCCCTGCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30173

QY 199 GTGGCACCGAGTCTCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 258
D  |||||
D 30172 GCGACGACCGCGCTGCAATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30113

QY 259 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 318
D  |||||
D 30112 GCGCGGTGCGCGCGCGCTCGAGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 30053

QY 319 GCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 378
D  |||||
D 30052 GCGCGCGCGAGCGACCGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 29993

QY 379 GCTCAGACTGTTTGGTAGCAACGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 438
D  |||||
D 29992 GCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 29933

QY 439 CGGCTCTTTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 498
D  |||||
D 29932 GACCGGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 29873

QY 499 AGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 558
D  |||||
D 29872 TGGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 29813

QY 559 CGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 618
D  |||||
D 29812 GCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 29754
```

```
QY 619 GAGCGGTGTATGGCGCGCGCTCACCATGTCGCTGAAGCCCGCGAGCAGCAGCAGCAG 678
D  |||||
D 29753 GAACGGCACACGACCGGGTGGCGAGCGCGCGCTTCTCGCGCGACACCGCGCGCAGCA 29694

QY 679 CAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 737
D  |||||
D 29693 CTGCAGCGCGCATCGAAGAGCGCGGATGAGCGCGAAGCCCTGCTCGGTGCGCGCGAC 29634

QY 738 TGCATAATGTCGCAAGCCCGCGCGCGCGCGCTTCTAGCTGTCGCGCGCGCGCGCGCTTC 797
D  |||||
D 29633 AGCGTCCGCGCAGCGCGCGCTGCGGTACAGCGTACAGGTCCGCGCGCGCAGAGCCCGCGAG 29574

QY 798 GCGTCTCTGTCCTCGGTCTCTCTGTCCTCGCGCGCGCGCTCCCTCTCGGTGCTCGCGC 857
D  |||||
D 29573 CCCCTGGAACCGCGCGCGCTGCGGTAGCCGTGCTCGACAGCTGCTGTTACCCCGCGCTC 29514

QY 858 GACCTCCGCGCGCGGAGCGCGCGCGCGCGCTGGGC 888
D  |||||
D 29513 GAGCGCGACCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 29483

RESULT 5
US-11-043-590-45
; Sequence 45, Application US/11043590
; Publication No. US20050277156A1
; GENERAL INFORMATION:
; APPLICANT: Compugen Ltd
; TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use therefor
; FILE REFERENCE: 1847.1011
; CURRENT APPLICATION NUMBER: US/11/043,590
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 45
; LENGTH: 1867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-043-590-45

Query Match 1.6%; Score 69.6; DB 7; Length 1867;
Best Local Similarity 77.8%; Pred. No. 2.1e-05;
Matches 84; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 648 GCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 707
D  |||||
D 439 GCAGCAGCAACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 498

QY 708 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 755
D  |||||
D 499 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 546

RESULT 6
US-11-137-671-13
; Sequence 13, Application US/11137671
; Publication No. US20050268350A1
; GENERAL INFORMATION:
; APPLICANT: Kirschbaum, Bernd
; APPLICANT: Berglund, Erick
; APPLICANT: Meisterernst, Michael
; APPLICANT: Polites, Greg
; TITLE OF INVENTION: PURIFICATION OF HIGHER ORDER TRANSCRIPTION COMPLEXES FROM TRANSGENIC NON-HUMAN ANIMALS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HELLER, EHRMAN, WHITE & McAULIFFE
; STREET: 1666 K Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```



Query Match	Best Local Similarity	Score	DB 7	Length	DB 6	Length	DB 5	Length	DB 4	Length	DB 3	Length	DB 2	Length	DB 1	Length	DB 0	Length	DB -1	Length	DB -2	Length	DB -3	Length	DB -4	Length	DB -5	Length	DB -6	Length	DB -7	Length	DB -8	Length	DB -9	Length	DB -10	Length	DB -11	Length	DB -12	Length	DB -13	Length	DB -14	Length	DB -15	Length	DB -16	Length	DB -17	Length	DB -18	Length	DB -19	Length	DB -20	Length	DB -21	Length	DB -22	Length	DB -23	Length	DB -24	Length	DB -25	Length	DB -26	Length	DB -27	Length	DB -28	Length	DB -29	Length	DB -30	Length	DB -31	Length	DB -32	Length	DB -33	Length	DB -34	Length	DB -35	Length	DB -36	Length	DB -37	Length	DB -38	Length	DB -39	Length	DB -40	Length	DB -41	Length	DB -42	Length	DB -43	Length	DB -44	Length	DB -45	Length	DB -46	Length	DB -47	Length	DB -48	Length	DB -49	Length	DB -50	Length	DB -51	Length	DB -52	Length	DB -53	Length	DB -54	Length	DB -55	Length	DB -56	Length	DB -57	Length	DB -58	Length	DB -59	Length	DB -60	Length	DB -61	Length	DB -62	Length	DB -63	Length	DB -64	Length	DB -65	Length	DB -66	Length	DB -67	Length	DB -68	Length	DB -69	Length	DB -70	Length	DB -71	Length	DB -72	Length	DB -73	Length	DB -74	Length	DB -75	Length	DB -76	Length	DB -77	Length	DB -78	Length	DB -79	Length	DB -80	Length	DB -81	Length	DB -82	Length	DB -83	Length	DB -84	Length	DB -85	Length	DB -86	Length	DB -87	Length	DB -88	Length	DB -89	Length	DB -90	Length	DB -91	Length	DB -92	Length	DB -93	Length	DB -94	Length	DB -95	Length	DB -96	Length	DB -97	Length	DB -98	Length	DB -99	Length	DB -100	Length	DB -101	Length	DB -102	Length	DB -103	Length	DB -104	Length	DB -105	Length	DB -106	Length	DB -107	Length	DB -108	Length	DB -109	Length	DB -110	Length	DB -111	Length	DB -112	Length	DB -113	Length	DB -114	Length	DB -115	Length	DB -116	Length	DB -117	Length	DB -118	Length	DB -119	Length	DB -120	Length	DB -121	Length	DB -122	Length	DB -123	Length	DB -124	Length	DB -125	Length	DB -126	Length	DB -127	Length	DB -128	Length	DB -129	Length	DB -130	Length	DB -131	Length	DB -132	Length	DB -133	Length	DB -134	Length	DB -135	Length	DB -136	Length	DB -137	Length	DB -138	Length	DB -139	Length	DB -140	Length	DB -141	Length	DB -142	Length	DB -143	Length	DB -144	Length	DB -145	Length	DB -146	Length	DB -147	Length	DB -148	Length	DB -149	Length	DB -150	Length	DB -151	Length	DB -152	Length	DB -153	Length	DB -154	Length	DB -155	Length	DB -156	Length	DB -157	Length	DB -158	Length	DB -159	Length	DB -160	Length	DB -161	Length	DB -162	Length	DB -163	Length	DB -164	Length	DB -165	Length	DB -166	Length	DB -167	Length	DB -168	Length	DB -169	Length	DB -170	Length	DB -171	Length	DB -172	Length	DB -173	Length	DB -174	Length	DB -175	Length	DB -176	Length	DB -177	Length	DB -178	Length	DB -179	Length	DB -180	Length	DB -181	Length	DB -182	Length	DB -183	Length	DB -184	Length	DB -185	Length	DB -186	Length	DB -187	Length	DB -188	Length	DB -189	Length	DB -190	Length	DB -191	Length	DB -192	Length	DB -193	Length	DB -194	Length	DB -195	Length	DB -196	Length	DB -197	Length	DB -198	Length	DB -199	Length	DB -200	Length	DB -201	Length	DB -202	Length	DB -203	Length	DB -204	Length	DB -205	Length	DB -206	Length	DB -207	Length	DB -208	Length	DB -209	Length	DB -210	Length	DB -211	Length	DB -212	Length	DB -213	Length	DB -214	Length	DB -215	Length	DB -216	Length	DB -217	Length	DB -218	Length	DB -219	Length	DB -220	Length	DB -221	Length	DB -222	Length	DB -223	Length	DB -224	Length	DB -225
-------------	-----------------------	-------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------	--------	---------



;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
;; TITLE OF INVENTION: DETECTION AND USES THEREOF

;; FILE REFERENCE: CL001559

;; CURRENT APPLICATION NUMBER: US/10/995,561

;; CURRENT FILING DATE: 2004-11-24

;; NUMBER OF SEQ ID NOS: 85702

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 13332

;; LENGTH: 94510

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (1)...(94510)

;; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-

US-10-995-561-13332

Query Match 1.5%; Score 68.4; DB 6; Length 94510;

Best Local Similarity 49.4%; Pred. No. 0.00023;

Matches 175; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 538 CCCTCCGCGCCGCTCCGCTCCGCGCGGTTCCGCGGCTCTCTTGGCGGCCCGGCTCCC 597

DB 6351 CCCTCCGCGCTCCGCGCTCCGCGCGGCTTTGTTGGTGGTAGGGGAGCAGAGCCG 6292

QY 598 GGCTGTCCCGCCGCGCTGCGAGCCGCTGTATGAGCCCTCACCATGTGCTGAAGCCC 657

DB 6291 AGTCAGAGACTTTTGCACGACCTTGGCGCGGAGCAGCGGATCAGCCGCTGCCGCGC 6232

QY 658 CACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 717

DB 6231 AAGATGCTGGAGCYGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 6172

QY 718 CAGCAGCGCGCGCGCGCGCTGCCAATGTC---CGCAAGCCCGCGCGCGCGCTTCTA 774

DB 6171 GAGAGGCGCGCTCGGGAGGCCATGCGGGCCCGGGGCTCCGGCCCGCGCCCGC 6112

QY 775 GCCTGCGCCCGCGCGCGCTTCGCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCG 834

DB 6111 GCTCCCG 6052

QY 835 GCTCCCT 888

DB 6051 GGGCGGCGCTCAGGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5998

RESULT 15

US-10-995-561-13381/c

;; Sequence 13381, Application US/10995561

;; Publication No. US2005072054A1

;; GENERAL INFORMATION:

;; APPLICANT: CARGILL, Michele et al.

;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

;; TITLE OF INVENTION: DETECTION AND USES THEREOF

;; FILE REFERENCE: CL001559

;; CURRENT APPLICATION NUMBER: US/10/995,561

;; CURRENT FILING DATE: 2004-11-24

;; NUMBER OF SEQ ID NOS: 85702

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 13381

;; LENGTH: 14248

;; TYPE: DNA

;; ORGANISM: Homo sapiens

US-10-995-561-13381

Query Match

Best Local Similarity 1.5%; Score 67.4; DB 6; Length 14248;

Matches 180; Conservative 0; Mismatches 156; Indels 9; Gaps 1;

QY 23 GCGCGCGCGCT 82

DB 5892 GCGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5833

QY 83 TCTGACCCCTCGACTTCGGTAAGAGTCCCTATCCGACCTCCGCTCCACCCGCGC 142

DB 5832 CGCTGCACCTGCTCTCTGCTCCACGCCACGGCGCTCCCTCCCGCCCCCTCCCGCCG 5773

QY 143 CTCGCGC-----GCCCGCCCTCCGATGCGCTCAGCGGCGCGCAGCTCTCTCGAGTC 193

DB 5772 CCGGGCGCTCGCGGCTGGCGGCTCTCCCGCTCTGTCTGTCCGTCGCTGGGTCCGCGGAG 5713

QY 194 CCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCCAGGTGGCCGCGGTGGCGCTCGC 253

DB 5712 CGACCTGGGCTAGGCAGGGGTCTAGGACGCGCGGAGGGGCGCAGCAGAGAGAACGG 5653

QY 254 TCCAGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 313

DB 5652 GCGCGGGGACCGCCAGGGGAGCCGACCTGCACCGGAGCGCAGGGGCGGGGCGGCGGAGC 5593

QY 314 CTTCCGCGCCCTCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 358

DB 5592 AAGCCCCCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 5548

Search completed: December 23, 2005, 06:22:29

Job time : 948 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame plus\_p2n model

Run on: December 23, 2005, 05:02:12 ; Search time 13935 Seconds  
(without alignments)  
5351.892 Million cell updates/sec

Title: US-10-802-228-2

Perfect score: 6961  
Sequence: 1 BLOSUM62.....TAHFPPYMTSPVQAHHQOOL 1312

Scoring table:  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n\_model -DSV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10802228/runat\_21122005\_110224\_9699/app\_query.fasta\_1.1479  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOCL=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45  
-DOCALIGN=200 -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10802228@cgn 1.10508 @runat\_21122005\_110224\_9699 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	6	AR153580 Sequence
2	6961	100.0	4481	6	AR401379 Sequence
3	6961	100.0	4481	6	AR447280 Sequence

4	6961	100.0	4481	6	AR632098	AR632098 Sequence
5	6961	100.0	4481	6	AX392465	AX392465 Sequence
6	6961	100.0	4481	8	HSU70323	U70323 Human ataxi
7	6890	99.0	4479	6	CQ727533	CQ727533 Sequence
8	6377	91.6	4163	8	HSDANSCA2	Y08262 H.sapiens m
9	6377	91.6	4200	6	A62706	A62706 Sequence 7
10	6025.5	86.6	4225	9	AF041472	AF041472 Mus muscu
11	5606.5	80.5	3798	6	AR632099	AR632099 Sequence
12	2612	37.5	2558	5	BC097692	BC097692 Xenopus 1
13	2380	34.2	2766	8	AK095017	AK095017 Homo sapi
14	1876	27.0	1257	6	AR447281	AR447281 Sequence
15	1876	27.0	1257	6	AR632110	AR632110 Sequence
16	1876	27.0	1257	6	AX392475	AX392475 Sequence
17	1876	27.0	1257	9	MMU70870	U70870 Mus musculu
18	1876	27.0	3781	8	HS3117973	AJ3117973 Homo sapi
19	1846	26.5	3758	8	HS3117974	AJ3117974 Homo sapi
20	1846	26.5	3893	8	HS3117972	AJ3117972 Homo sapi
21	1846	26.5	3897	8	HS3117971	AJ3117971 Homo sapi
22	1846	26.5	4380	8	HS3117970	AJ3117970 Homo sapi
23	1832	26.3	4283	8	BC068012	BC068012 Homo sapi
24	1819	26.1	3635	8	BC082760	BC082760 Homo sapi
25	1819	26.1	3764	9	BC054483	BC054483 Mus muscu
26	1816.5	26.1	4674	8	AF034373	AF034373 Homo sapi
27	1780.5	25.6	3920	8	AY188335	AY188335 Homo sapi
28	1775.5	25.5	3980	8	AY188337	AY188337 Homo sapi
29	1773.5	25.5	4272	8	AY188334	AY188334 Homo sapi
30	1768.5	25.4	4012	8	AY188336	AY188336 Homo sapi
31	1768.5	25.4	4496	8	AY188338	AY188338 Homo sapi
32	1481	21.3	1080	9	MMZ78162	Z78162 M.musculus
33	1426	20.5	168247	8	AC005014	AC005014 Homo sapi
34	1310	18.8	3486	9	BC043451	BC043451 Mus muscu
35	1305	18.7	3570	5	BC057536	BC057536 Danio rer
36	1273	18.3	107717	8	AC137055	AC137055 Homo sapi
37	1257	18.1	4369	6	CQ851373	CQ851373 Sequence
38	1257	18.1	4369	8	AK128613	AK128613 Homo sapi
39	1245.5	17.9	108139	14	AC004085_3	Continuation (4 of
40	1222	17.6	2809	9	BC034083	BC034083 Mus muscu
41	1151	16.5	182703	14	AC073117	AC073117 Homo sapi
42	1012.5	14.5	623	6	AR159546	AR159546 Sequence
43	907	13.0	572	6	AR159558	AR159558 Sequence
44	857	12.3	516	6	AR447279	AR447279 Sequence
45	857	12.3	516	6	AR632097	AR632097 Sequence

#### ALIGNMENTS

RESULT 1  
LOCUS AR153580 AR153580 4481 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 18 from patent US 6235872.  
ACCESSION AR153580  
VERSION AR153580.1 GI:15121112  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4481)  
AUTHORS Breiden, D.E. and Rabizadeh, S.  
TITLE Proapoptotic peptides dependence polypeptides and methods of use  
JOURNAL Patent: US 6235872-A 18 22-MAY-2001;  
FEATURES Location/Qualifiers  
source  
1..4481  
/organism="unknown"  
/mol\_type="unassigned DNA"

#### ORIGIN

Alignment Scores: 2.26e-97 Length: 4481  
Pred. No.: 6961.00 Matches: 1312  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB:

```
US-10-802-228-2 (1-1312) x AR153580 (1-4481)
Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGCGCTCAGCGCGCGCAGCTCTCGAGTCCCGGCTGGCCACCGAGTCTCGCGCTTC 222
Qy 21 AlaAlaAlaArgTrrProGlyTrrArgSerLeuGlnArgProAlaAlaArgSerGlyArg 40
Db 223 GCCGAGCGCAGGTGGCGCGGCTCGCTCCAGCGCGCGCGCGCGAGCGCGGCGG 282
Qy 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 283 GCGCGCGGTGGCGCGCGCGCGGACGATCCCTCCGCGCGCGCTCCCGCGCGCGCGCC 342
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 343 GCGCGCGCTCCCTCCCGCGCAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAAC 402
Qy 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgLeuLeuGlyGlyGlyPro 100
Db 403 GCGAAACGCGCGCGCGCGCTTTCGCGCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGCGCT 462
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAla 120
Db 463 CCGCGCGCTTCGTGCTGCTCTTCCTCCCTCCGCGCGCGCGCGCGCGCTCCCGCGCGG 522
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 523 CCNACCGCGCTCCCGCTCGCGCGCGGTGGTCCCGCGCGGTTCGCGCGCTCTCTTG 582
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 583 GCGCGCGCGCTCCCGCTCCCGCGCGGTGGTCCCGCGCGGTTCGCGCGCTCTCTTG 642
Qy 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
Qy 181 GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly 200
Db 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
Db 763 AGCGGCTTCTAGCGTCGCGCGCGCGCGCTTCGCGCGCTTCGCTCGCTCTCTCTCT 822
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
Db 823 TCCTCGCGCACGGCTCCCTCTCGTGGTTCGCGCGACCTCCGCGCGCGCGCGCGCGCG 882
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db 883 CTGGCGAGGCTCGAAACAGTAACAGACTCGCTCAGCTACAGATTTCTTTTATGGA 942
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db 943 ATCTATGCAAAATATGAGATGTTTATATATCATCATCATGTTTGGTTCCTCAATGTGAA 1002
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 1003 GTACAAAGTGAAAAATGGAGGTATATGAAGAGATTTTAAACACTTACAGTCCGAGGTG 1062
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db 1063 GATTTGGTACTTGTATGCCCATCATGAAAGTACAGAAATCCAGTTCGGGCGCGGAAACGT 1122
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db 1123 GAAGAAATAATGAGAGATTTTGTTCAAATGTTTTCAGACTTTTGTGTGTGTCAGTTTAAA 1182
Qy 341 AspMetAspSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db 1182 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1242
Qy 361 ValAsnGlyGlyHisLysGlyLysAspLeuGluProTrrAspAlaGlyGlyLeuThrAla 380
Db 1243 GTGAATGGCGAACAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1302
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrrAspProAsnAspMet 400
Db 1303 AATGAGAACTTGAGGCTTTGGAAATGACGATCTAATGATGGATGCCCAATGATG 1362
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValSerThrTyrAspSerSerSerSerSer 420
Db 1363 TTTTGATATATGAGAAATAATGAGTGTAGTGTCTACGATATGATAGAGTTTATCTTCG 1422
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1423 TATACAGTGCCTTAGAAAGAGATACTCAGAGAAATTTTAAACGGGAAGCAAGGCA 1482
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACAGTGTAGCAGAGAAATTTAGTCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 1542
Qy 461 AsnAspAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATGATGATGAGTGGAGAGAAATAACACAGCTTACAGAGAAATTTCCAGTGAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db 1603 GAGGGCACAGCATAAACACTAGGGAAAAATAATATATATCTCTCCGCAAAAGAAATAGA 1662
Qy 501 GluValIleSerTrrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTCGGGAAAGTGGAGAGAGAAATTTACCGCGGTATGGCGCAGCTTGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACATTTCCAGATTTCAACCCGAAATCTGGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyGlyValProTrrProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCCATCTCTCTCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCACCTTCTCGCTACAGTCCAGTCCCAACTCTCTTCCACCTCGCGCAGCACCCCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACAGCGCGCGCTCCAGCGCGCTTCGCGCGCATCCAGACCCCGTCTCACCCCTCTGCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGTTCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2022
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 2023 AGGATGTCCCAAGGCGCGCAGCATCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT 2082
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 2083 TCCATATCCAGTGGCTTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCT 2142
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTrrSerSerValSerGlyValPro 680
Db 2143 CCAGTAGCAGGACCATCTCCGGGGAAACGTGGTCACTCAGTGGTCACTGGGCTTCCA 2202
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTATCCCTAAACCTCATAGACCCAGGTCTCCACAGACAGAAACAGTATTTGGAATACC 2262
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db 2263 CCCAGTGGCGCAGTCTTGTGCTTCTCCCAAGCTGGTATTTTCCAACCTGAAGCTGTGCC 2322
```

```
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCGTATTCCAGCTGCATCTCTACGCGCTGTAGTCTCTGCATCGAACAGAGCTGTACC 2382
QY 741 ProSerSerGluAlaIysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCTGAGGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAGAAAATATTAAACCCATGAACATCACCTAGTCTCTCAAAAGCTGAAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATCACCACTGTTCTGAACATAGAAAACAGATTGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTTACAGCCAGTTCTACTTCTGAATCTATGGATCAACTACTATAACAAA 2622
QY 821 AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGAGAAAATCAAGAGATTGTATCAAGACAAAATTTGAACCAAGTGTAG 2682
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTTCTTCAATGAAATAGCAGCAGCACTGTACAGTGGCAGCAGCAAGCCGAATAG 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisArgGlyProGluValThr 880
DB 2743 CCAGCAATTTCCCTCTTCAATACTTAGTAACACGAGCACAAGAGGGAGCCTTGAGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnLysAspLysGluGlu 900
DB 2803 TCCCAAGGGTTCCAGCTTCCAGCCAGCATGTAAACAGAGAAAACGATAGGAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrIleuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGACGACGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGGAGTTT 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB 2923 AACCCAGTTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCGCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
DB 2983 GCACAACTTAGCCCATCTATGGTGGGTCAACAAGCCAACTCCAGTTTATCTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTTTGCACCAATATGATGATCCAGTCCAGTGAGCCAGCGCGTGCACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
DB 3103 TACCAATACCTATGACGCCCATGCCAGTCAATCAAGCCAAAGACATATAGCAGTACCA 3162
QY 1001 AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCCAACAGCGGCAAGACCATCATCAGATGCCATGATGACCCAGCGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
DB 3223 GCAGCGGCCCCAGGATTCAGGACCCACCCAGCTTACTCCAGCAATATGTTGCCATC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
DB 3283 AGTCCTCAGCAGTTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATATATCATCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCCTCATGTCTATAGTCTGTGAATACAGGGTAATGCTAGAATGATGGCACCACCAACA 3402
```

```
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
DB 3403 CACGCCCCAGCGTGGTTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGCG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
DB 3463 CATGCGATGTATGCATGCTCCCAAAATACCATACAACAGGAGACAAGCCCTCTTCTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTGCAATTTCCAGCGGCTCCCTTGTCTCAGCAGTATGCGCACCTAACCTACCTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCTCAGCTTCCAGTACCTCCCTTGGACAGCAGCAAGCCCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
DB 3643 GTTGAAGTCACTCTGACCCAGTCTCTGTTCAGCACCATCAGCACCGCCGCCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTACACGCGGGCTTGGCCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTCTCCATGACACCTGCTCCCAACACGAGTCGCCAGCAATAGTTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
DB 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGTTCAGCCGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGCCCACTACCTCAGGCTCATGTACAGTCAGTCAGGAATGTTCTTCTCATTCACA 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaLeu 1280
DB 3943 GCCCATGCCCCAATGATGCTAATGACGACACAGCCACCCGCGGGTCCCAGGCGGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCATCAGGCCCATTCAGTCTCGACAACAGCGCATTTCCCTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
DB 4063 CACCCTTCAGTACAGCCCAAGCCCAACCAACAGCAGTTG 4098
```

## RESULT 2

AR401379 4481 bp DNA linear PAT 18-DEC-2003  
LOCUS Sequence 20 from patent US 6623927.

DEFINITION AR401379

ACCESSION AR401379

VERSION AR401379.1 GI:40148692

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 4481)

AUTHORS Brahmachari,S.K., Choudhry,S., Mukerji,M. and Jain,S.

TITLE Method of detection of allelic variants of SCA2 gene

JOURNAL Patent: US 6623927-A 20 23-SEP-2003;

Council of Scientific and Industrial Research; New Delhi;

INX; Location/Qualifiers

source 1..4481

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores: 2.26e-97 Length: 4481

Pred. No.: 6961.00 Matches: 1312

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-802-228-2 (1-1312) x AR401379 (1-4481)			
QY	1	MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe	20
Db	163	ATGCGCTCAGCGCGCGAGCTCTCGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTC	222
QY	21	AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerglyArg	40
Db	223	GCCGACGCGAGGTGGCGCGGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGGCGG	282
QY	41	GlyGlyGlyGlyAlaAlaProGlyTyrProSerAlaAlaProProProGlyPro	60
Db	283	GGCGGGGTGGCGCGCGCCCGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGGCC	342
QY	61	GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn	80
Db	343	GGCCCCCTCCCTCCCGCGCAGAGCTCGCTCCCTCCGCTCAGACTGTGTTTGGTAGCAAC	402
QY	81	GlyAenGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro	100
Db	403	GGCAACGGCGCGCGCGTTCGCGCCCGGCTCCCGCGGGCTCTCTGGTCTCGCGCGGCT	462
QY	101	ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla	120
Db	463	CCCCGCCCTTCGTCTGCTCTCTCCCTCCCTCCGACGCGCGCGCGCCCTCCGCGCGCG	522
QY	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu	140
Db	523	CCAACCGCGCCTCCCGCTCGCGCGCGTGGCTCCCGCGCGTTCGCGCGTCTCCTTG	582
QY	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160
Db	583	GGCGCGCGCGCTCCCGCGGTGTCGCGCGCGGTGCGAGCGGGTGTATGGCGCCCTCACC	642
QY	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180
Db	643	ATGTCGTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702
QY	181	GlnGlnGlnGlnGlnGlnGlnProProProAlaAlaAlaValArgLysProGlyGly	200
Db	703	CAGCAGCAGCAGCAGCAGCAGCGCGCGCGCGCGCTGCCAATGTCCGCAAGCCCGCGCGC	762
QY	201	SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerSerSer	220
Db	763	ACGCGCCTTCTAGCGCTCGCGCGCGCGCGCTTCGCGCTCTCGCTCTCGTCTCCTCG	822
QY	221	SerSerAlaThrAlaProSerSerValValAlaAlaAlaThrSerGlyGlyArgProGly	240
Db	823	TCCCTCGGCCACGGCTCCCTCTCGGTGTGCGCGGACCTCCGCGCGGAGCGCGCGCG	882
QY	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260
Db	883	CTGGCAGAGGTGCAAAACAGTAACAAAGGACTGGCTCAGTCTACGATTTCTTTTGTGGA	942
QY	261	IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280
Db	943	ATCTATGCAAAATATGAGGATGTTATATGAGGAGTGTATTAACACTTACAGTCCGGAAGTGT	1002
QY	281	ValGlnValLysAenGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300
Db	1003	GTACAAAGTCAAAATATGAGGATGTTATATGAGGAGTGTATTAACACTTACAGTCCGGAAGTGT	1062
QY	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerSerGlyProLysArg	320
Db	1063	GATTTGGTACTTGATCCGACATGAGAAAGTACAGATCCAGTTCGCGGCGCGCAACGT	1122
QY	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys	340

1123	GAAGAAATAATCGAGAGTATTTTGTTCAAATGTTTCAGACTTGTGTGGTACAGTTTAAA	1182	
QY	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
Db	1183	GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA	1242
QY	361	ValAenGlyGluHisLysGluLysAspLeuGluProTyrTrpAspAlaGlyGluLeuThrAla	380
Db	1243	GTGAATGGCGAAACAAAGAGAGAGCCTGGAGCCTGGATGCGAGTGAACCTACAGACC	1302
QY	381	AenGluGluLeuGluAlaLeuGluAsnAspValSerAenGlyTyrTrpAspProAsnAspMet	400
Db	1303	AATGAGGAACCTTGAGGCTTTGGAAATGACGTATCTAATGATGGATGCCAATGATATG	1362
QY	401	PheArgTyrAenGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420
Db	1363	TTTCGATATATGAAGAAAATTTATGGTGTAGTGTCTACGTATGATGAGCAGTTTATCTTCG	1422
QY	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440
Db	1423	TATACAGTCCCTTAGAAAAGAGATAACTCAGAAGAAATTTTAAAAACGGGAAGCA	1482
QY	441	AenGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460
Db	1483	AACCAAGTTAGCAGAGAAGAAATTTGAGTCAAGTCCCAAGTACAAAGCTCGAGTGGCCCTG	1542
QY	461	AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg	480
Db	1543	AATGATGATAGGAGTGGAGAGAGAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACT	1602
QY	481	GluGlyHisSerIleAenThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500
Db	1603	GAGGGCAGACATAAACACTAGGGAAAAATAATATATATCTCTCTCGACAAAGAAATAGA	1662
QY	501	GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520
Db	1663	GAAGTCATATCTCGGGGAAGTGGGAGACAGAAATTCACCGGTATGCGGCCAGCCTGATCG	1722
QY	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540
Db	1723	GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGGTTCA	1782
QY	541	AspGlnArgValValAsnGlyGlyValProTyrTrpProSerProCysProSerProSerSer	560
Db	1783	GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCTCTCTCT	1842
QY	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro	580
Db	1843	CGCCACCTTCTCGCTACAGTCAGTCCCAACTCTCTTCCACCTCGGCGACGCCCTCT	1902
QY	581	ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla	600
Db	1903	ACACGGCGCCCTCCAGGCCCTTCGCGGCCATCCAGACCCCGCTCTCACCCCTCTGCT	1962
QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
Db	1963	CATGGTCTCCAGTCTCTGCTCTACTATGCTTAAACGCATGTCTTCAGAGGGCTCCA	2022
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
Db	2023	AGGATGTCCTCCAAAGGCCCGACATCTCTGAAATCACAGAGTTTCTGCTGGAGGGGT	2082
QY	641	SerIleSerSerGlyLeuGluPheValSerHisAenProProSerSerGluAlaAlaThrPro	660
Db	2083	TCCATATCCAGTGGCCTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCTCT	2142
QY	661	ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro	680
Db	2143	CCAGTAGCAAGACAGTCCCTCGGGGGGAACGTGTCATCAGTGTGTCAGTGGGTTTCCA	2202
QY	681	ArgLeuSerProLysThrHisArgProArgSerProArgGlnAenSerIleGlyAenThr	700
Db	2203	AGATTATCCCTTAAACTCATAGACCCAGGTCTCCGACAGAACAGTATTGGAAATACC	2262

```

QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720
DB 2263 CCCAGTGGGCCAGTCTTCTCCAGCTGGTATTTATTTCCAACTGAAGCTGTGGCC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCCCTATTTCCAGCTGCATCTCTCAGCTGCTGCTAGTCTGCATCGNACAGAGCTGTACC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CCTTCTAGTAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCTGCAGGG 2442
QY 761 AsnLysGluAsnLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAAGAAAATATTAAACCAATGAACATCACCTAGTCTTCTCAAAAGCTGAACACAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATACACAGTGTGTTCTGNACATAGAAAACAGATTGATGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGTTACAGCAAGTCTCTCTCTGAATCTATGATCAACTACTAAACAAA 2622
QY 821 AsnArgGluGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGGAGAAAATCAAGAGATTGATCAAAAGACAAAATTTGAACCAAGTGTCTAAG 2682
QY 841 AspSerPheIleGluAsnSerSerAsnGlyThrSerGlySerSerLysProAsnSer 860
DB 2683 GATCTCTTATGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCGGAATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB 2743 CCAGCATTTCCCTTCAATCTAGTACACGAGCACAAGAGGGGACCTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
DB 2803 TCCCAAGGGGTTCCAGACTTCCAGCCCGCAGCATGTAAACAGAGAAAGACGATGAAGGAAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGAGCGCAGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGGTTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB 2923 AACCCAGTTCCTTCTCTCAGCAAAAGCCTTCTACTACCCCACTTCACTCGGCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
DB 2983 GCACAACTAGCCCATCTATGTGGGTCTATCAACAGCCCACTCCAGTTTATATCTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTGTTGACCAATATGATGTATCCAGTCCCAAGTGGCCAGCGCGTGCAACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
DB 3103 TACCCATATCTATGACGCCATCCAGTGAATCAAGCCNAGACATATAGAGCAGTACCA 3162
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCCAACAGCGGCAAGACCAAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
DB 3223 GCAGCGGGCCACCGATTGAGCGCCACCCCAAGCTTACTCCAGCAATATGTGTGCTAC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
DB 3283 AGTCCTCAGCAGTTTCCAAATAGCCCCCTTTGTTACGATGTGCCACATATCATGTCTCAG 3342

```

```

QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCTCATGTCTTAGTCTCTGTAATACAGGTAATGCTAGTAATGATGGCACCACAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
DB 3403 CACGCCCAAGCTGGTTTAGTATCTTCTTTCAGCAACTCAGTACGGGGCTCATCAGCAGACG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
DB 3463 CATGGGATGTATGTCATGTCCCAAAATTACCATACAAACAAGGAGACAGGCCCTCTTTCTTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTGCCATTTCCACGGGTCTCTCTCAGCAGTATGGCACCTTAACGTTACCTGTCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCTCAGCCTTTCAGTACCCCTCAGGACAGCAGCAAGCCAAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
DB 3643 GGTGGAAGTCTATCTCTGACCCAGTCTCTGTTTCAGCACCATCAGCAGCCGCCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCCATTTTACCACGGGGCTTCGGCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTCTCATGACACCTTCCCTCCAAACGCGAGTCCGACAGAAATAGTTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
DB 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGTTCAGCCGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGGCCCATCTACCTCAGGCTCATGTACAGTTCAGGAATGGTTCCTTCTCATCCAAC 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaLeu 1280
DB 3943 GCCCATGCGCCCAATGATGCTAATGACGACACAGCCACCCGCGGTCCCAGGCCGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCACTACAGCCCATTCAGTCTCGAACACAGCGCATTTCCCTTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312
DB 4063 CACCCTTCAGTCAAGCCCAACCCACCAACAGCAGTTG 4098

```

RESULT 3  
 LOCUS AR447280 4481 bp mRNA linear PAT 20-FEB-2004  
 DEFINITION Sequence 2 from patent US 6673535.  
 ACCESSION AR447280  
 VERSION AR447280.1 GI:42675575  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 4481)  
 AUTHORS Pulst,S.M.  
 TITLE Methods of detecting spinocerebellar Ataxia-2 nucleic acids  
 JOURNAL Patent: US 6673535-A 2 06-JAN-2004;  
 FEATURES Cedars-Sinai Medical Center; Los Angeles, CA  
 1. .4481  
 Location/Qualifiers  
 /organism="unknown"  
 /mol\_type="mRNA"  
 source  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,26e-97 Length: 4481  
Score: 6361.00 Matches: 1312  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-802-228-2 (1-1312) x AR447280 (1-4481)

Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
Db 163 ATGCGCTCAGCGCGCGAGCTCCTCGGAGTCCGCGGTGGCCACCGAGTCTCGCGCTTC 222  
Qy 21 AlaAlaAlaArgTrrProGlyTrrArgSerLeuGlnArgProAlaArgSerGlyArg 40  
Db 223 GCCCGAGCCAGGTGGCGCGGTGGCGTCCGTCACAGCGCGCGCGCGAGCGGGCGG 282  
Qy 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60  
Db 283 GCGCGCGGTGGCGCGCGCGCGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCC 342  
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80  
Db 343 GCGCGCGCTCCCTCCCGCGAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 402  
Qy 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100  
Db 403 GCGCAACGCGCGCGCGGTTCGCGCGCGCTCCCGCGGTCTCTTGGTCTCGCGCGCGCT 462  
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120  
Db 463 CCGCGCGCGCTTCGTCGTCTCTTCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCG 522  
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140  
Db 523 CCAACCGCGCGCTCCCGCGCTCGCGCGCGGTGGTCTCCCGCGCGGTTCGCGGTCTCTTG 582  
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
Db 583 GCGCGCGCGGTTCGCGGTGTCCCGCGCGCGGTGCGAGCGGTGTATGGCGCGCTCACC 642  
Qy 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
Db 643 ATGTCGTGAAGCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702  
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
Db 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762  
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220  
Db 763 AGCGCGCTTCTAGCGTCGCGCGCGCGCGCTTCGCGCTTCGCTCGGTCTCCTCG 822  
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240  
Db 823 TCCTCGCGCACGGCTCCCTCCTCGGTGGTGGCGGAGCTCCGATGTCGCGCGCGCGCG 882  
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
Db 883 CTGGGCAGAGGTGGAACAGTAAACAGAGCTGCTCAGTCTACGATTTCTTTTATGGA 942  
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
Db 943 ATCTATGCAATATCAGGATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGA 1002  
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
Db 1003 GTACAAGTGAAGAAATGGAAGGTATATATGAAGAGTATTTTAAACCTTACAGTCCGAAGTGT 1062  
Qy 301 AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg 320  
Db 1063 GATTTGGTACTTGATCCGCGCATGATGAGAAAGTACAGAAATCCAGTTTCGGGCGCGAAGCT 1122

Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys 340  
Db 1123 GAAGAAATAATGGAGAGTATTTTGTTCAAATGTTTCAGACTTGTGTGTGTCAGTAAAA 1182  
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
Db 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTCTAAA 1242  
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrrAspAlaGlyGluLeuThrAla 380  
Db 1243 GTGAATGCGGAACACAAAGAGAGAGCTCGAGCGCTGGGATGCGAGTGAATCAGACCC 1302  
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrrAspProAsnAspMet 400  
Db 1303 AATGAGGAATCTGAGGCTTTGGAAAAATGACGTATCTAATGATGGATGCCCAATGATATG 1362  
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTrrAspSerSerLeuSerSer 420  
Db 1363 TTTGATATAATGAAGAAAAATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
Db 1423 TATACAGTGCCTTTAGAAAGAGATAACTCAGAAAGATTTTTTAAACCGGAAGCAAGGCA 1482  
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTrrLysAlaArgValAlaLeuGlu 460  
Db 1483 AACCACTTAGCAGAAGAAATTTGAGTCAAGTCCCGCAGTACAAAGCTCGAGTGGCCCTGAA 1542  
Qy 461 AsnAspAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
Db 1543 AATGATGATAGAGGTGAGGAAGAAAAATACACAGCAGTTCAGAGAAATTTCCAGTGAACGT 1602  
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTrrIleProProGlyGlnArgAsnArg 500  
Db 1603 GAGGGGCACAGCATAAACACTAGGGAAAAATAAATATATTCCTCTCCGCAAGAAATAGA 1662  
Qy 501 GluValIleSerTrrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
Db 1663 GAAGTCATATCCTGGGAAGTGGGAGACAGAAATTTACCGGTATGGCGCAGCTGGATCG 1722  
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACATTTCTCAGATTTCAACCCGAAATCTTGGTTC 1782  
Qy 541 AspGlnArgValValAsnGlyValProTrrProSerProCysProSerProSerSerSer 560  
Db 1783 GACCAAGAGTAGTAAATGGAGGTGTTCCCTGGCCATCGCCCTTCCCAATCTCTCTCTCT 1842  
Qy 561 ArgProProSerArgTrrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
Db 1843 CGCCACCTTCTCGCTACCACTACAGTCAAGTCCCAACTCTCTTCCACTCGGGCAGCCACCT 1902  
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
Db 1903 ACACGGCGCGCTCCAGCGCGCGCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Db 1963 CATGGTCTCCAGTCTCTGCTCTACTATGCTTAAACGATGTCTTTCAGAAAGGCGCTCCA 2022  
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Db 2023 AGGATGTCCTCCAAAGGCCACGACATCTCTGAAATCACAGATTTCTGTGGGAGGGGT 2082  
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
Db 2083 TCCATATCCAGTGGCTTAGAATTTGATCCACAAACCCACCCAGTGAAGCAGTACTCTCT 2142  
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTrrSerSerValValSerGlyValPro 680  
Db 2143 CCAGTAGCAGACAGTCCCTCCGGGGGAACGTGGTTCATCAGTGGTTCAGTGGGTTCCA 2202

QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
DB 2203 AGATTATCCCTAAACCTCATAGACCAGGTCTCCAGACAGAAAGTATTGGAAATACC 2262  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
DB 2263 CCAGTGGGCCAGTCTCTGCTCTCTCCCAAGCTGGTATTATTCCTCAACTGNACTGTGTGCC 2322  
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
DB 2323 ATGCCTATTCCAGCTGCATCTCTACGCCCTGCTAGTCTGCTGCATCGAACAGAGCTGTATTACC 2382  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAlaGlnArgGlnAsnSerProAlaGly 760  
DB 2383 CCTTCTAGTGGAGCTAAAGATTCCAGGCTTCAGATCAGAGGACAGAACTCTCTCTGCAGGG 2442  
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
DB 2443 AATAAAGAAATATTAAACCCATGAACATCACCTAGTCTTCTCAAAAGCTGAACAA 2502  
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspLeuLysLysPheLys 800  
DB 2503 GGTATATCACCACTGCTTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2562  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
DB 2563 AATGATTTTAGGTTACAGGCAAGTCTTACTTCTGAATCTATGATCAACTACTAAACAA 2622  
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
DB 2623 AATAGAGAGGGAGAAATCAAGAGATTGATCAAGACAAATTTGACCAAGTGTCTAAG 2682  
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
DB 2683 GATTCCTTCAATGAATAGCAGCAGCACTGTACCAAGTGGCAGCAGCAAGCCGATAGC 2742  
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
DB 2743 CCCAGCATTTCCCTCTCAATCTTAGTAACACAGCAGCACAAGAGGGGAGCTCTGAGGTCACT 2802  
QY 881 SerGlnGlyValGlnThrSerProAlaCysLysGlnGluLysAspLysGluGlu 900  
DB 2803 TCCCAAGGGGTTCCAGACTTCCAGCCAGCATGTAAACAGAGAAACCATAGGAAGAG 2862  
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
DB 2863 AAGAAAGACGACCTGAGCAGTTAGGAATCAACTGATCCCAATGCCAAGGAGTTC 2922  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrProThrSerProArgProGln 940  
DB 2923 AACCCAGTTCCTCTCTCAGCCAAAGCCCTTCTACTACCCCAACTTCACCTCGGCCCTCA 2982  
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnPro 960  
DB 2983 GCACCAACCTAGCCCATCTATGGTGGGTCATCAACAGCCCACTCCAGTTATATCTCAGCCT 3042  
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
DB 3043 GTTTGTTTGCACCAATATGATGTATCCAGTCCAGTGAGCCAGCGGTGCAACCTTTA 3102  
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
DB 3103 TACCCCAATACCTATGACGCCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162  
QY 1001 AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020  
DB 3163 AATATGCCCCAACAGCGGCAAGACCAACATCATCAGAGTGCCATGATGACCCAGCGTCA 3222  
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
DB 3223 GCAGCGGGCCCGATGAGCCACCCACAGCTTACTCCAGGCATATGTTGCTTAC 3282  
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060

DB 3283 AGTCTCAGCAGTGTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAG 3342  
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
DB 3343 CATCTCTATGCTATAGTCTGTAAATACAGGGTAAATGTAGAAATGATGGCACCAACA 3402  
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
DB 3403 CACGCCCAAGCCCTGGTTAGTATCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGAG 3462  
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
DB 3463 CATGGATGTATGCTATGCTGCTCCAAATTTACCATCAACAAGGAGACAAGCCCTTCTTCTAC 3522  
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
DB 3523 TTTGCCATTTCCACGGGCTCCCTTCTCAGCAGTATGCGACCTTAAGCTACCTCGCAC 3582  
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160  
DB 3583 CCACATATCTCCACACCTCAGCTTTCAGTACCTCCCTCAGCAGTATGCGACCTTAAGCTACCTCGCAC 3642  
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaGlnAla 1180  
DB 3643 GGTGAAGTCACTCTGCACTCCAGTCTGTTTACGACCATCAGCAGCAGCCGCCAGGCT 3702  
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTACCTTACCACGGGGCTTCGCCCAACT 3762  
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
DB 3763 CCACCTCTCATCAGCAGCTTCCAAACAGCAGTCCGCCACAGAATAGTTTCCACAGCAGCA 3822  
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
DB 3823 CAACAGACTGTCTTACGATCCATCTCTTCTCAGCTTACCGCGGTATACCAACCCACCC 3882  
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
DB 3883 CACATGGCCACGTCACCTCAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCCAAT 3942  
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaLeu 1280  
DB 3943 GCCCATGCGCCAATGATGCTAATGACACACAGCCACCCCGGGGTCCCGAGGCCCTC 4002  
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
DB 4003 GGTCAAAGTGCATCAGCCCATTCAGTCTCGAACACAGCGCATTTCCCTATATGACG 4062  
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312  
DB 4063 CACCTTTCAGTACAGCCCAACCAACAGCAGTTG 4098

## RESULT 4

ARG32098  
LOCUS 4481 bp mRNA linear PAT 14-FEB-2005  
DEFINITION Sequence 2 from patent US 6844431.  
ACCESSION AR632098  
VERSION AR632098.1 GI:59773824  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
AUTHORS 1 (bases 1 to 4481)  
TITLE Nucleic acid encoding spinocerebellar ataxia-2 and products related thereto  
JOURNAL Patent: US 6844431-A 2 18-JAN-2005;  
FEATURES  
source Cedars-Sinai Medical Center; Los Angeles, CA  
Location/Qualifiers  
1. .4481



/organism="unknown"									
/mol_type="mRNA"									
ORIGIN									
Alignment Scores:									
Pred. No.:	2.26e-97	Length:	4481						
Score:	6961.00	Matches:	1312						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	6	Gaps:	0						
US-10-802-228-2 (1-1312) x AR632098 (1-4481)									
Qy	1	MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe	20						
Db	163	ATGCGCTCAGCGCGCGAGCTCCTCGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTC	222						
Qy	21	AlaAlaAlaArgTrrProGlyTrrArgSerLeuGlnArgProAlaArgSerGlyArg	40						
Db	223	GCCGAGCAGGTGGCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGGCGG	282						
Qy	41	GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro	60						
Db	283	GGCGCGGTGGCGGGCGCCCGGACCGATATCCCTCCGCGCGCCCTCCCGCGCGCC	342						
Qy	61	GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn	80						
Db	343	GGCCCCCTCCTCCCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC	402						
Qy	81	GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyPro	100						
Db	403	GGCAACGGCGCGCGCTTCGCGCGCGCTCCCGCGGCTCTTGGTCTCGCGCGGCT	462						
Qy	101	ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla	120						
Db	463	CCCCCGCCCTTCGTGCTGCTCTCTTCGCCCTCGCGCAGCGCGGGCGCCCTCCGCGCGG	522						
Qy	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu	140						
Db	523	CCAAACCGCGCTCCCGCGTGGCGCGCGTCCCGCGCGCTTCGGCGCTCTCTCTTG	582						
Qy	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160						
Db	583	GCGCGCGCGCTCCCGCGTGTCCCGCGCGCGTGGAGCGGTGTATGGCGCCCTCAC	642						
Qy	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180						
Db	643	ATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702						
Qy	181	GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly	200						
Db	703	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	762						
Qy	201	SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer	220						
Db	763	AGCGGCTTCTAGCGTCGCGCGCGCGCGCTTCGCGCTCTCTCGTCTCTCGTCTCG	822						
Qy	221	SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly	240						
Db	823	TCCTCGGCCACGGCTCCCTCCTCGTGGTTCGGCGCGACCTCCGGCGGGGAGGCGCGG	882						
Qy	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260						
Db	883	CTGGCGAGAGTTCGAAACAGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTGATGGA	942						
Qy	261	IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280						
Db	943	ATCTATGCAAAATATGAGGATGGTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAA	1002						
Qy	281	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300						
Db	1003	GTACAAGTGAAATATGAGGTATATATGAAGGAGTTTTTAAACTTACAGTCCGAGTGT	1062						

Qy	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg	320
Db	1063	GATTTGGTACTTGGTGGCGCATGAGAAAGTACAGAATCCAGTTCGGGGCCGAAACGT	1122
Qy	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys	340
Db	1123	GAAGAAATAATGGAGAGTATTTTGTTCAAATGTTCAGACTTGTGTGTGACAGTTTAAA	1182
Qy	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
Db	1183	GATATGAGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA	1242
Qy	361	ValAsnGlyGluHisLysGluLysAspLeuGluProTrrAspAlaGlyGluLeuThrAla	380
Db	1243	GTGAATGGCGAAACAAAGAGAGGACCTGGAGCCCTGGGATGCAGGTGAACACTCACAGCC	1302
Qy	381	AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrrAspProAsnAspMet	400
Db	1303	AATGAGGAACCTTGAGGCTTTTGAAAAATCACGTATCTAATGGATGGATCCCAATGATATG	1362
Qy	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420
Db	1363	TTTCGATATAATGAAGAAAAATTATGGTGTAGTGTCTACGTATGATGACAGTTTATCTTCG	1422
Qy	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440
Db	1423	TATACAGTGCCCTTAGAAAGAGATAAAGCTCAGAAGAAATTTTTTAAACGGGAAGCAAGGCA	1482
Qy	441	AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460
Db	1483	AACAGATTAGCAGAAAGAAATTTAGTCAAGTGCACAGTACAAAGCTCGAGTGGCCCTGGA	1542
Qy	461	AsnAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg	480
Db	1543	AATGATGATAGGAGTGGAGAGAAAAATACACAGCAGTTTCAGAGAAATTCACAGTGAACGT	1602
Qy	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500
Db	1603	GAGGGCACAGCATAAACACACTAGGAAAAATAATATATATCTCTCTGGACAAAGAAATAGA	1662
Qy	501	GluValIleSerTrrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520
Db	1663	GAACTCATATCTCGGGGAAGTGGAGACAGAAATTCACCGGTATGGGCGACGCTGGATCG	1722
Qy	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540
Db	1723	GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCCGAATTCCTGGTTCA	1782
Qy	541	AspGlnArgValValAsnGlyGlyValProTrrProSerProCysProSerProSerSer	560
Db	1783	GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCATCGCCTTGCCCATCTCTCTCTCT	1842
Qy	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro	580
Db	1843	CGCCACCTTCTCGCTACCACTAGTCCCAACTCTCTTCACCTCGGCGGACCCCTCT	1902
Qy	581	ThrArgProProSerArgProProSerArgProProSerHisProSerSerAla	600
Db	1903	ACACGGCGCCCTCCAGGCCCTTCGGGGCATTCAGACCCCTCTCACCCCTCTCTCT	1962
Qy	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
Db	1963	CATGGTCTCCAGCTCCTGTCTCTACTATGCTAAAGCATGTCTTCAGAAAGGCGCTCCA	2022
Qy	621	ArgMetSerProLysAlaGlnArgHisProArgHisArgValSerAlaGlyArgGly	640
Db	2023	AGGATGTCCCAAGGGCCCGACATCCCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT	2082
Qy	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro	660
Db	2083	TCCATATCCAGTGGCGCTAGAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCT	2142

```
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
DB 2143 CCAGTAGCAAGACCAAGTCCCTCGGGGGAACGTGTGTCATCAGTGTGCTAGTGGGTTCAC 2202
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
DB 2203 AGATTATCCCTTAAAACTCATAGACCCAGGTCTCCAGACAGAACAGTATTGGAAATACC 2262
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
DB 2263 CCCAGTGGCCAGTCTTCTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGCC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCCTATTCCAGCTCATCTCCTACGCTGTAGTCTCTGTCATCGAACAGAGCTGTATTAC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCGAGNACTCTCTCTGAGGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAGAAATATTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATCACCACTGTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTATACAGCAAGTCTTACTTCTGAATCTATGGATCAACTACTTAAACAAA 2622
QY 821 AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGAGAAATCAAGAGATTGATCAAGACAAAATTGAAACNAGTGTAG 2682
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTCCTTCAATGAAATAGCAGCAGCAACTGTATCCAGTGGCAGCAGCAAGCCGAATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisArgGlyProGluValThr 880
DB 2743 CCAGCATTTCCCTTCAATCTTAGTAACACCGAGCACAAGAGGGGACCTTAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGlyLysAspLysGluGlu 900
DB 2803 TCCCAAGGGTTCAGACTTCACCCCGCAGCATGTAAACAGAGAAAGCAGTAAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAATGCCAATGCCAAGGAGTTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB 2923 AACCCAGTTCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCCTCA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
DB 2983 GCACAACCTAGCCCATCTAGTGGGTCACTCAACAGCCCACTCAAGTTTATCTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTTTTCACCAATATGATGTATCCAGTCCAGTGAGCCAGCGGTGCAACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
DB 3103 TACCCAAATACCTATGACGCCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCA 3162
QY 1001 AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCCAACAGCGGCAAGACCAACATCATCAGAGTGCCATGATGCAACCCAGGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProAlaTyrSerThrGlnTyrValAlaTyr 1040
```

```
DB 3223 GCAGGGGGCCCAAGATTGCAGGCCACCCACAGCTTACTCCACCAATATGTTGCCCTAC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
DB 3283 AGTCTCTAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCTCTATGTCTATAGTCTCTGTAATACAGGGTAAATGCTAGAAATGATGCAACCAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyValHisGlnGlnThr 1100
DB 3403 CACGCCCAAGCCCTGGTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
DB 3463 CATGCGATGTATGTCATGTCCTCCAAATTTACCATACAAGGAGACAAGCCCTTCTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTTGCATTTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCTAACGCTACCTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCCCTCAGCTTTCAGTACCCCTCAGTACGAGCAGCAAGCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180
DB 3643 GGTGGAAGTCTCTCTGACCCAGTCTGTTCAGCACCATCAGCACCAGGCCGCCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCACTTTACACGCGGGCTTGCCTCACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTCTCATGACACCTGCTCCCAACAGCAGTCCGCAAGTATAGTATTCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
DB 3823 CAACGACGTCTTCTTACGATCCATCTCTCAGTTCAGCCGGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGCCCGTACCTCAGGCTCATGTATCAGTACAGTACAGGAATGTTCTTCTCATCCA 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
DB 3943 GCCCATGCGCCCAATGATGCTAATGACGACACAGCCACCCCGGGTCCCCAGGCCGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCATCAGCCCATTCAGTCTCGACAAACAGCGCATTTCCCTTATATGAGC 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
DB 4063 CACCCTCTAGTACAGCCCAAGCCCAACCAACAGCAGCTTG 4098
```

## RESULT 5

```
AX392465 4481 bp DNA linear PAT 23-MAR-2002
LOCUS Sequence 1 from Patent WO0216417.
DEFINITION AX392465
ACCESSION AX392465
VERSION AX392465.1 GI:19700764
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Pulst,S.M. and Huynh,D.P.
```

TITLE	Transgenic animal model for neurodegenerative disease and uses thereof		
JOURNAL	Patient: WO 0216417-A 1 28-FEB-2002; CEDARS-SINAI MEDICAL CENTER (US)		
FEATURES	Location/Qualifiers		
source	1..4481		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
CDS	163..4101		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD29012.1"		
	/db_xref="GI:19700765"		
	/translation="NR5AAAAPSPAPVATERRFAAARPGWRSLQRPARRSGRGGG AAGPYPAAPPPGPPGPPQSPGASDFGSGNGGCGAFRPGSRRLGLGPPPR PVPVLLPLASGAPPAATRASPLGARASPPRGVSLARPACGPPRACPEPVVGLPT MSLKPPQQQQQQQQQQQQQQQQQQQQPPPAANVRKPGGGLLASPAAPASPSSSV SSSAFSSVVAATSGGGRGLGRGNKGLPOSTISFDGIYANMMRWHLITSVVG SKCEVQVNGGIYEVFKTYGPKCDLVLDAAHEKSTESSGPKKEEIMESILPKCSDF VVOGKDMSSYAKRDAFTDSAISAKVNGEHKEKDLDPWAGELTANELEALENDVS NGWDMDFRYNEENYGVSTVDSLSSTVPLERDNEEFLEKREARMANOLAEIES AQYKARVALENDRESEKTYTAVOENSSERECHSINTRENKVIIPGORNREVIISWGSC RONSRMQPGSGSPSSTSHTSDFNENSGSDQRVNCGVPWSPCPSPSPSPSRY QSGNPLPRAATITRPPSRPPSPRPPSPHPSAHGSPAPVSTMPKRNSSGPPRMSPP KAQRHPRHRVSRAGSISLGGLEFVSHNPPSEAAITPPVARTSPSGTWSVSVGVPRL SPKTRHPRSPRGISNTPPSGPLASPOAGIIPTEAVAMPIPAASPTPASPASNAVIT PSSEAKDRLDORONSPAGNKENIKNETSPSPKSNKGISPVVSHRKOIIDLKX KPNDFRLOPSTSEMDOLLANKREGEKSRDLIKDKIEPSAKDFSFIENSNCCTSGSS FNKSPSISPSILNTEHRRGPEVTSQGVQTSPPACKQSKDKKDKDAEOVRKSTLN PNAKEFPNSYQKPSITTPSPQAQPSPMVGHQOQTPVYTPQVCFAPNNMYFPV VSPGVQPLYPIMTPMNVNOAKTYRAVPMFQORQDQHQSMHMPASAAAGPPAATP PAYSTQYVAYSQQPFPNQLVQHPHYQSPHYVSPVIOGNARMABPHTHAQPLP SSATQYGAHEOTHAMVACPILPKYNKETSFPFVAILSTGSLAQOAHNPATLHPHTPH QPSATPTGQOQSHGSGPSPVQHHOHAQALHLASPOQOQSAIYHAGLAPTPSM TPASNTQSPQSPFPAQOQTVFTIHPSHVQPAYTNPPMAHVPQAHVQSGWVPSHTPH APMMLMTTQPPGGPQAALQALQIPVPTTATFPFYMTHPSVQAHHQOOL"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,26e-97	Length:	4481
Score:	6961.00	Matches:	1312
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-802-228-2 (1-1312) x AX392465 (1-4481)			
QY	1	MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe	20
DB	163	ATGCGCTCAGCGCGCGCAGCTCCTCGAGTCCGCGGTGGCCACCGAGTCTCGCGCTTC	222
QY	21	AlaAlaAlaArgTTPProGlyTTPArgSerLeuGlnArgProAlaArgSerClyArg	40
DB	223	GCCGAGCAGGTGCGCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGG	282
QY	41	GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro	60
DB	283	GCGCGCGGTGGCGCGCGCGCGCGCGATCCCTCGCGCGCGCGCGCGCGCGCGCG	342
QY	61	GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn	80
DB	343	GGCCCCCTCCCTCCCGCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTTGGTAGCAAC	402
QY	81	GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro	100
DB	403	GGCAACGGCGCGCGCGGTTCGCGCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGCG	462
QY	101	ProArgProPheValValValLeuLeuProLeuAlaSerProGlyValaProProAlaAla	120
DB	463	CCCCGCCCTTCGTGCTGCTCTTCTCCCCCTCGCAGCCCGCGCGCGCGCGCGCG	522
QY	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerIeu	140
DB	523	CAAACCGCGCTCCCGCTCGCGCGCGTGCCTCCCGCGCGTTCGGCGCTCTCCTTG	582
QY	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160
DB	583	GCGCGCGCGCTCCCGCTGTCCTCCCGCGCGCGTGCAGCGCGGTATGGGCGCTCACC	642
QY	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180
DB	643	ATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702
QY	181	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
DB	703	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	762
QY	201	SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerSerSer	220
DB	763	AGCGGCTTCTAGGCTCGCGCGCGCGCGCGCTTCCGCGCTCCTCGCTCTCCTCG	822
QY	221	SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly	240
DB	823	TCCTCGGCCAGGCTCCTCTCTCGTGGTCCGCGCGACCTCCGCGCGCGCGCGCGCG	882
QY	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260
DB	883	CTGGCAGAGGTGGAACAGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTGATGA	942
QY	261	IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280
DB	943	ATCTATGCAATATGAGGATGGTTCATATCTTACATCAGTTGTGGCTCCAAATGTGAA	1002
QY	281	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300
DB	1003	GTACAAGTGAATAATGAGGTATATATGAAGGAGTTTAAAACTTACAGTCCCAAGTGT	1062
QY	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg	320
DB	1063	GATTTGGTACTTGTATGCGGCACATGAGAAAGTACAGAAATCCAGTTCGGGCGCAACCT	1122
QY	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys	340
DB	1123	GAAGAAATAATGAGAGATATTTTGTTCAAATGTTTCAGACTTTGTGTGTACAGTTTAAA	1182
QY	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
DB	1183	GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA	1242
QY	361	ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla	380
DB	1243	GTGAATGCGCAACACAAAGAGAGAGACCTGGAGCGCTGGGATGCAGGTGAATCAGACCC	1302
QY	381	AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTTPAspProAsnAspMet	400
DB	1303	AATCAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATG	1362
QY	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420
DB	1363	TTTCGATATATGAAGAAATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1422
QY	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440
DB	1423	TATACAGTCCCTTTAGAAAGAGATAAATCTCAGAAAGATTTTAAACCGGAGCAAGGCA	1482
QY	441	AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460
DB	1483	AACAGTTAGCAGAAAGAAATTTAGTCAAGTCCCGCAGTACAAAGCTCAGTGGCGCTGAA	1542
QY	461	AsnAspAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg	480
DB	1543	AATGATGATGAGGTGAGGAAGAAATAACACACAGTTTCAGAGAAATTCACAGTGAACGT	1602
QY	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500

Db 1603 GAGGGGCAAGCATAAACACTAGGAAATAATAATATATTCCTCCGGCAAGAAATAGA 1662  
Qy 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
Db 1663 GAAGTCATATCTCTGGGAAGTGGAGACAGAATTACCGCGTATGGCCAGCGCTGGATCG 1722  
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCACCCGGAATTCCTGGTTCA 1782  
Qy 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560  
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCTCGCCATCGCCTTGGCCATCTCTCTCTCTCT 1842  
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
Db 1843 CGCCCACTTCTCGCTACCGTCCAGTCCCACTCTCTTCCACTCGGCGCCGCCACCCCT 1902  
Qy 581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600  
Db 1903 ACAGCGCCGCCCTCCAGGCCCCCTCGCGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Db 1963 CATGGTCTCCAGCTCTCTCTACTATGCTTAAACGCGATGCTTTCAAGAGGGCTCCA 2022  
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Db 2023 AGGATGTCCCCAAGGCCCGCAGCAGCATCTCGNAATCACAGAGTTTCTCTGGAGGGT 2082  
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAlaAlaThrPro 660  
Db 2083 TCCATATCCAGTGGCTAGAAATTGTATCCCAACCCAGTCCAGTCAAGCAGCTACTCTCT 2142  
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTipSerSerValValSerGlyValPro 680  
Db 2143 CCAGTAGCAAGGACCAAGTCCCTCGGGGGAAACGTGTGTATCATAGTGTGAGTGGGTCCA 2202  
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
Db 2203 AGATTATCCCTTAAACTCATAGCCAGGTCTCCAGACAGAACAGTATGGAAATACC 2262  
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
Db 2263 CCCAGTGGCCAGTCTCTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGCC 2322  
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
Db 2323 ATGCCTATTCCAGTGCATCTCTTACGCCTGTAGTCTCTGCATCGAACAGAGCTGTACC 2382  
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
Db 2383 CCTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCTGAGGG 2442  
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
Db 2443 AATAAGAAATAATTAACCAATGAACATCATCTACCTAGCTTCTCMAAAGCTGAACACAA 2502  
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysPheLys 800  
Db 2503 GGTATATCACAGTGTGTTCTGAACATAGAAAAACAGATTGATTTAAAGAAATTTAAG 2562  
Qy 801 AsnAspPheArgLeuGlnProSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
Db 2563 AATGATTATAGTTTACAGCCAAAGTTCTACTTCTGAATCTATGGATCAACTACTTAACCAA 2622  
Qy 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
Db 2623 AATAGAGGGGAGAAAAAATCAAGAGATTGTATCAAGACAAAAATTTGAACCAAGTGTAA 2682  
Qy 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860

Db 2683 GATTCTTTTCATTGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCCAATAGC 2742  
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
Db 2743 CCCAGCATTTCCCTTCAATACTTAGTAACACGGAGCAACAAGAGGGGACCTTGAGGTCACT 2802  
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900  
Db 2803 TCCCAGGGGTTTCAGACTTCAGCCAGCATGTAAACAAGAGAAACAGATAGAAGAGAG 2862  
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
Db 2863 AAGAAAGACGAGCTGACCAAGTTAGGAAATCAACATTTGAATCCCAATGCCAAGAGATT 2922  
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
Db 2923 AACCCAGCTTCTCTCTCAGCCAAAGGCTTCTACTACCCCAACTTCACTCGGCTCAA 2982  
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
Db 2983 GCACAACTTAGCCCATCTATGGTGGGTCTATCAACAGCCCACTCCAGTTTATACTCAGCT 3042  
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
Db 3043 GTTCTGTTTGCACCAATATGATGTATCCAGTCCCACTGAGCCAGGCGTGCAACCTTTA 3102  
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
Db 3103 TACCAATACCTTATGACGCCCATGCGCAGTGAATCAAGCCCAAGACATATAGACGAT 3162  
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020  
Db 3163 AATATGCCCCAAACAGCGGCAAGCAGCATCATCAGAGTCCCATGATGACCCAGCTCA 3222  
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
Db 3223 GCAGCGGGCCCACTGACGCCCACTTGCAGCCCACTTACTCCACGCAATATGTTGCCTAC 3282  
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
Db 3283 AGTCCTCAGCAGTCTCCCAATCAGGCCCTTGTTCAGCATGTGCCATTTATCAGTCTCAG 3342  
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
Db 3343 CATCTCTATGTCTATAGTCTGTAAATACAGGGTAATGCTAGATGATGACCAACCAACA 3402  
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
Db 3403 CAGCCCGAGCTGTTTGTAGTATCTTCTTCCAGCAACTCAGTACGGGCTCATGAGCAGCG 3462  
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
Db 3463 CATGCGATGTATGATGTCTCCCAATTTACCATACAAAGAGAGCAAGCCCTCTCTTCTAC 3522  
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
Db 3523 TTTGCCATTTCCACGGGCTCCCTTGTCTCAGCAGTATGCCACCCTAACCGTACCTGCGAC 3582  
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160  
Db 3583 CCACATCTCCACACCTCAGCCTTCCAGTACCCCTTCCAGTCCAGCAGCAGCAAGCAACAT 3642  
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla 1180  
Db 3643 GGTGGAAGTCACTCTGCACCCAGTCTGTGTTTTCAGCAGCATCAGCAGCCAGGCGCCAGGCT 3702  
Qy 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
Db 3703 CTCATCTGGCCAGTCCACAGCAGCAGTCCAGCATTTTACCACCGGGGCTTGGCCCACT 3762  
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
Db 3763 CCACCTTCCATGACACTGCTGCTCCACACGCGAGTCCGCCAGAAATAGTTTCCCGCAGCA 3822

```
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
D 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGCTTCAGCGCGGTATATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
D 3883 CACATGGCCCGACCTCAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyGlyProGlnAlaAlaLeu 1280
D 3943 GCCCATGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
D 4003 GCTCAAGTGCCTACAGCCCATTCAGTCTCGAACAAGCGCAATTCCTCCCTATATGAGC 4062
QY 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
D 4063 CACCCTTCAGTACAGCCCAACCAACACGACGATGTTG 4098
RESULT 6
HSU70323
LOCUS HSU70323 4481 bp mRNA linear PRI 20-NOV-1996
DEFINITION Human ataxin-2 (SCA2) mRNA, complete cds.
ACCESSION U70323
VERSION U70323.1 GI:1679683
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 4481)
Pulat.S.-M., Nechiporuk,A., Nechiporuk,T., Gispert,S., Chen,X.-N.,
Lopes-Gendes,I., Pearlman,S., Starkman,S., Orozco-Diaz,G.,
Lunkes,A., DeJong,P., Rouleau,G.A., Auburger,G., Korenberg,J.R.,
Figueroa,C. and Sahba,S.
Moderate expansion of a normally biallelic trinucleotide repeat in
spinocerebellar ataxia type 2
Nat. Genet. 14 (3), 269-276 (1996)
8896555
2 (bases 1 to 4481)
Pulat,S.-M.
Direct Submission
Submitted (10-SEP-1996) Medicine, Cedars-Sinai, 8700 Beverly Blvd.,
Los Angeles, CA 90048, USA
FEATURES
Location/Qualifiers
1..4481
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24.1"
1..4481
/gene="SCA2"
163..4101
/gene="SCA2"
/standard_name="spinocerebellar ataxia type 2"
/codon_start=1
/product="ataxin-2"
/protein_id="AAB19200.1"
/db_xref="GI:1679684"
/tranlation="MRSAAAPRSPAVATERRFAARWPGWRSLQRRPARRSGRGGG
AAGPYPSAAPPFGPPPPSPASDFCGSNGNGGAFRPGSRLLGLGGPPR
PFVVLLPLASPGAPPAATRASPLGAPSPRSVSLARAPCPRPACFPVGLPT
MSLKQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
SSSATAPSSVVAATGGGRLGGRNSKGLPOSTISFDGIYANRMVHLTSVVG
SKCEVQXNGIYEGFTYSPKCDLVLDAAHEKSTESSSGKREEMESILFKCSDF
VVQPKDMSSYAKRDVFTDSAIKAVNGEHKEKLEPDAGELTANELEALENDVS
NGWDPDMFRYNEENYGVSTVDSLSSTYVPLERDNEEFLKREARANQIAERTSS
AQYKARVALENDRESEKTYAVORNSEREGHSINTRENYIPPGQRNREIVSGSG
RQNSPRMQPGSGNSPFRSTSHTSDFNPNSSDQKRVNNGVFWPSPSPSPSPSRY
```

```
OSGNSLUPPRAATPTPRSPRPPSPRPPSHPSAHGSPAPVSTWPKRMSSGPPRMSP
KAQHPNRHVRVSAGRGSISSGLEFVSHNPSEAAATPPVARTSPSGGTWSSVVGPRLL
SPKTHRPRSQRNSIGNTPSGPVLASPOAGIIPTEAVAMPIPAASPVSHPASRAVTP
PSSAKDSRLQDORNSPAGNKENIKNETSPSKAENKIGISPVSEHRKQIDDLKK
FKNDPRLOPSSSTESMDQLLNKREGEKSRDLIKDKLEPSAKDSFIENSSNCTSGSS
KNPSFSPSILSNTEHRGVEVTSQGVOTSSPACQKQKDEKKAQAEVRSKSTLN
PNAKEFNPRSPSPKPTTTPSPQAPSPSMVGHQOQTPVYTPVCFAFMNMYFVP
VSPGQLYLPMTFMPVNAQTYRAVPNMQDQHQHQSAMMHHPASAAGPPTAATP
PAYSTQVAVSPQOFPNQLVQHVPYVSPVIOGNRMAAPTHAQGLVPS
SSATQGAHEQTHAMAYACPUPYNKETSPSYFAISTGSLAQOYVAPNATLHPHTPH
OPSATPTGCGSQHSGSHPPAPSPVQHGHQAALHLASPOQQAISYAGLAPTPPSM
TPASNTQSPQNSFAAQOTVFTIHPSHVQPAATNPMAHVPOAHVQSGMVPSPHTAH
APMLMTTQPPGPGQAALASALQPIPVSTTTHAFPYMTHPSVQAHHQOOL"
ORIGIN
Alignment Scores: 2.26e-97 Length: 4481
Pred. No.: 1312
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-802-228-2 (1-1312) x HSU70323 (1-4481)
QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
D 163 ATGCGCTCAGCGCGCGAGCTCCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTC 222
QY 21 AlaAlaAlaArgTTPProGlyTTPArgSerLeuGlnArgProAlaArgSerGlyArg 40
D 223 GCCGAGCAGCGTGGCGCGGTGGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCG 282
QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
D 283 GCGCGCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 342
QY 61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
D 343 GCGCGCGCTCCTCCCGCAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAlaGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyGlyPro 100
D 403 GCGAACGCGCGCGCGCGCGTTCGCGCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGG 462
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyValAlaProProAla 120
D 463 CCGCGCGCTTCTGCTGCTGCTCTTCTCCCTCCCGCAGCGCGCGCGCGCGCGCGCG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
D 523 CCAACCGCGCGCTCCCGCTCGCGCGCGCTCGCTCCCGCGCGCTTCCCGCGCTCTCTTC 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
D 583 GCGCGCGCGCTCCCGCTCGCGCGCTCCCGCGCGCTCGCGCGCTGATGGCGCGCTCAAC 642
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
D 643 ATGCGCTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
D 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerSerSer 220
D 763 AGCGCGCTTCTAGCGTCCCGCGCGCTTCCCGCGCTTCCCGCTTCTCTCTCTCTCT 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240
D 823 TCCTCGCGCACCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 882
```

QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
DB 883 CTGGGCGAGAGGTGGAACAGTAAACAAAGAGCTGCCTCAGTCTACGATTTCTTTTGATGA 942  
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
DB 943 ATCTATGCAATATAGAGTGGTTTCATATACATCATGTTGCTGCCAATGTGAA 1002  
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
DB 1003 GTACAAAGTGAATAATGAGGTATATATGAAGAGTATTTTAAACCTTACAGTCGGAAGTGT 1062  
QY 301 AspLeuValLeuAspAlaHisGlyLysSerThrGluSerSerGlyProLysArg 320  
DB 1063 GATTTGGTACTTGTATCCGCACATGAGAAAGTACAGAAATCCAGTTTCGGGGCGGAACGT 1122  
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
DB 1123 GAAGAAATATAGAGATGATTTTGTTCAAATGTTTCAGACTTGTGTGTACAGTTTAA 1182  
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
DB 1183 GATATGAGCTCCAGTTATGCAAAAGAGAGATGCTTTTACTGACTCTGCTATCAGTGTAAA 1242  
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380  
DB 1243 GTGAATGGCGAAACACAAAGAGAGACCTGGAGCCCTGGGATGCAAGTGAACCTCACAGCC 1302  
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
DB 1303 AATGAGGAACCTTGAGGCTTTGGAAATGAGCTATCTAATGGATGGGATCCCAATGATATG 1362  
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
DB 1363 TTTTCGATATATAGAGAAATTTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
QY 421 TyrThrValProLeuGluArgAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
DB 1423 TATACAGTGCCTTAGAAAGAGATAACTCAGAAAGAAATTTTAAACCGGAAGCAAGGGCA 1482  
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
DB 1483 AACCGATTAGCAGAGAAATTTAGTCAAGTGCCTGAGTACAAAGCTCGAGTGGCCCTGGAA 1542  
QY 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
DB 1543 AATGATGATAGGAGTGGAGAGAAATAACACAGCAGTTCAGAGAAATTCACGTGAACGT 1602  
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
DB 1603 GAGGGGCACAGCATAAACATAGGGAATAATAATATATTTCTCTCTGACACAAAGAAATAGA 1662  
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
DB 1663 GAAGTCAATATCTCTGGGAACTGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCG 1722  
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
DB 1723 GGCTCCATGCTATCAAGATCACTTCTACACTTCAGATTTCAACCCGAATTCGTGTCA 1782  
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560  
DB 1783 GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTTCCCTCTCTCTCT 1842  
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
DB 1843 CGCCACCTTCTGCTACCAAGTCAAGTCCCAACTCTCTCCACTCTCGGGCAGCACCCCT 1902  
QY 581 ThrArgProProSerArgProProSerArgProProSerHisProSerAla 600  
DB 1903 ACAGGGCGGCTCCAGGCCCCCTCGGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620

DB 1963 CATGTTTCTCCAGCTCTCTGTCTACTATGCTAAACGATGTCTTTCAGAGGGCTCCA 2022  
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
DB 2023 AGGATGTCCCCAAAGGCCAGCGACATCTCGAAATCACAGAGTTTCTGCTGGAGGGGT 2082  
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
DB 2083 TCCATATCAGTGGCTAGAAATTTGTATCCCAACACCACCAGTGAACGACTACTCTCT 2142  
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680  
DB 2143 CCAGTAGCAAGACAGTCCCTCGGGGGAACCTGCTCATCAGTGGTCACTGGGTTC 2202  
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
DB 2203 AGATTTATCCCCATAAATCATAGACCAGTCTCCAGACAGACAGATATTGGAATATCC 2262  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
DB 2263 CCAGTGGGCCAGTTCTTGTCTTCTCCCAAGCTGTTATTTTCCAACCTGTAAGCTTGGC 2322  
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
DB 2323 ATGCTTATTTCCAGCTGCATCTCTTACGCTGCTAGTCTGTCATCGAACAGAGCTGTACC 2382  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
DB 2383 CCTTCTAGTAGGCTAAAGATTTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGG 2442  
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
DB 2443 AATAAAGAAATATTAAACCCAATGAACCATCACCTAGTCTCTCAAAAGCTGAAACAA 2502  
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
DB 2503 GGTATATACCAAGTTGTTTCTGAACATAGAAACAGATTTGATTTAAAGAAATTTAAG 2562  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
DB 2563 AATGATTTTAGGTTACAGCCAAGTTCTACTTCTGAATCTATGATCAACTACTATAACAA 2622  
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
DB 2623 AATAGAGAGGAGAAAAATCAAGAGATTTGATCAAAAGACAAAAATTTGAACCAAGTGTAG 2682  
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
DB 2683 GATTTCTTTCATTTGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAGCGAATAGC 2742  
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
DB 2743 CCCAGCATTTCCCTTCAATACTTAGTAACCGAGCACAAAGAGGGGACCTGAGGTCAC 2802  
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900  
DB 2803 TCCCAAGGGGTTCCAGACTTCCAGCCCGCATGTAAACAAAGAGAAACGATAAGGAAGAG 2862  
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
DB 2863 AAGAAAGAGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCCCAATGCAAGGAGTTC 2922  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
DB 2923 AACCCAGTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCCTCAA 2982  
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
DB 2983 GCACAACTTAGCCCATCTATGGTGGGTCTCAACAGCCAACTCCAGTTTATCTACGCT 3042  
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980

Db 3043 GTTGTGTTGGACCAATATGATGATATCCAGTCCAGTGAGCCCGAGGGGTGCAACCTTTA 3102  
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
Db 3103 TACCCAATACCTATGAGCGCCCATGCAGTGAATCAAGCAACATATAGAGCAGTACCA 3162  
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020  
Db 3163 AATATGCCCCAAACAGCGGCAACAGCAGCATCATCAGAGTGCCATGATGCACCGCTCA 3222  
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
Db 3223 GCAGGGGGCCCAACGATTCAGCCACCCACAGCTTACTCCAGCATATGTTGCCATAC 3282  
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
Db 3283 AGTCCTCAGCAGTCCCAAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 3342  
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaAlaArgMetMetAlaProProThr 1080  
Db 3343 CATCCTCATGTCTATAGTCTGTAATACAGGGTAATGCTAGAAATGATGGCACCAACA 3402  
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
Db 3403 CACGCCCAAGCCTGGTITAGTATCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3462  
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
Db 3463 CATGCGATGATGTCATGCTCCCAAAATACCATACACAGGAGACAGGCCCTTCTTCTAC 3522  
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
Db 3523 TTTGCAATTTCCACGGGCTCCCTTGTCTCAGCAGTATGGCACCTTAACGCTACCCCTGCAC 3582  
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160  
Db 3583 CCACATCTCCACACCCCTCAGCTTTCAGTACCCCACTGGACAGCAGCAAGCCAAACAT 3642  
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180  
Db 3643 GGTGAAGTATCTCTGCACCCAGTCTCTTTCAGCACCATCAGCACCGCGCCAGCGCT 3702  
Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCATTTTACACGCGGGGCTTTCGCCAACT 3762  
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
Db 3763 CCACCTTCATGACACCTGCCTCCAAACAGCAGTCCGCACAGAAATAGTTTCCAGCAGCA 3822  
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
Db 3823 CAAACAGACTGTCTTTACGATCCATCTCTCAGCTTCAGCGCGGTATACCAACCCACCC 3882  
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
Db 3883 CACATGGGCCCACTACCTCAGGCTCATGTACAGTCAGGAATGGTTCTCTCATCCAACT 3942  
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
Db 3943 GCCCATGGCCCAATGATGCTAATAGCAGACAGCCACCGCGGGTCCCGAGCGCCCTC 4002  
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
Db 4003 GCTCAAGTGCACTACAGCCCAATTCCAGTCTCGACAAACAGCGCAATTTCCCTATATGAG 4062  
Qy 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312  
Db 4063 CACCCTTCAGTACAAGGCCCAACCAACAGCAGTGTG 4098

RESULT 7  
CQ727533 4479 bp DNA linear PAT 03-FEB-2004  
. LOCUS CQ727533

DEFINITION Sequence 13467 from Patent WO02068579.  
ACCESSION CQ727533  
VERSION CQ727533.1 GI:42293778  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1  
REFERENCE Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
AUTHORS Kits, such as nucleic acid arrays, comprising a majority of  
TITLE humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 13467 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES Location/Qualifiers  
source  
1..4479  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,65e-96 Length: 4479  
Score: 6890.00 Matches: 1308  
Percent Similarity: 99.77% Conservative: 1  
Best Local Similarity: 99.70% Mismatches: 3  
Query Match: 98.98% Indels: 2  
DB: 6 Gaps: 0  
US-10-802-228-2 (1-1312) x CQ727533 (1-4479)  
Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
Db 163 ATGCGCTCAGCGCGCGAGCTCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGG-TTC 221  
Qy 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgArgSerGlyArg 40  
Db 222 GCCGACAGCAGTGTGCCGGGTGGCGTCTCGCTCCAGCGCGCGCGCGCGCGCGCG 281  
Qy 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60  
Db 282 GCGCGCGTGGCGGGCGCGGACCGTATCCCTCCGCGCCCTTCCTCCCGCGCGCGCC 341  
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80  
Db 342 GGCCCCCTCCCTCCCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTGTGTAGCAAC 401  
Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100  
Db 402 GGCAACGCGCGCGCGGCTTTCGGCCCGGCTCCCGCGGCTCCTTGGTCTCGCGCGGCT 461  
Qy 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120  
Db 462 CCGCGCCCTTCGTCT 521  
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140  
Db 522 CCAACCGCGGCTCCCGCTCGCGCGCGCGCTCCCGCGCGCTCCCGGCTCTCTCTTTG 581  
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
Db 582 GCGCGCGCGCTCCCGGCTGTCCCGCGCGCGCGGTGCGAGCCGGTGTATGGGCCCTCAC 641  
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
Db 642 ATGTCGCTGAAGCCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 701  
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
Db 702 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 761



QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerValSerSer 220  
DB 762 AGCGGCTTCTAGCGTCGCGCGCGCGCT - TCGCGCTCCTCGCTCCTCG 820  
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240  
DB 821 TCTCGCGCACGGCTCCCTCTCGGGTTCGGGGACCTCCGGCGCGGGAGGCCCGGC 880  
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
DB 881 CTGGGCGAGGTCGAACAGTAACAAGGACTGCCTCAGCTACGATTCTTTTGATGGA 940  
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
DB 941 ATCTATGCAAAATATGAGGATGGTTATATATACATTCAGTTGTGGCTCCAAATGTGAA 1000  
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
DB 1001 GTACAAAGTGAAAAATGGAGTATATATGAAGGATTTTTTAAACCTTACAGTCCGAAGTGT 1060  
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320  
DB 1061 GATTTGGTACTTGCATCCGCACATGAGAAAGTACAGATCCAGTTCCGGGCGGAAACGT 1120  
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
DB 1121 GAAGAAATAATGGAGAGTATTTGGTTCAATGTTACAGCTTTGTGTGTACAGTTTAA 1180  
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
DB 1181 GATATGGACTCCAGTTTATGCAAAAGAGATGCTTTTACTGACTCTCTATCATAGTGTCTAAA 1240  
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProIleAspAlaGlyGluLeuThrAla 380  
DB 1241 GTGAATGGCGAACCAAGAGAGGACCTGGAGCCCTGGGATCGAGTGAATCTCACAGCC 1300  
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
DB 1301 AATGAGGAACTTCAGGCTTTGGAAATGACGTATCTAATGGATGGATCCCAATGATATG 1360  
QY 401 PheArgTyrAsnGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
DB 1361 TTTTCGATATAATGAAGAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1420  
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
DB 1421 TATACAGTGCCTTAGAAGAGATTAATCTAGAGAAATTTTAAACCGGAGCAGGGCA 1480  
QY 441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
DB 1481 AACCAGTTAGCAGAGAAATTTAGTCAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAA 1540  
QY 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
DB 1541 AATGATGATAGGAGTGGAGGAGAAATATACACAGCAGTTCAGAGAAATTTCCAGTGAACGT 1600  
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
DB 1601 GAGGGGCACAGCATAAACACTAGGAGAAATAAATATATTTCTCTGTGACAAAGAAATAGA 1560  
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
DB 1661 GAAGTCATATCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCG 1720  
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
DB 1721 GGCTCCATGCGCATCAAGATCCACTTCTCACACTTTCAGATTCAACCCGAAATTCGTGTCA 1780  
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560  
DB 1781 GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCCCTTGGCCCATCTCTCTCTCT 1840  
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580

DB 1841 CGCCACACTTCTCGCTACAGTCCCACTCTTTCCACCTCGGGCAGCCACCCCT 1900  
QY 581 ThrArgProProSerArgProSerArgProSerArgProProSerHisProSerAla 600  
DB 1901 ACACGGCGCGCTCCAGGCCCCCTCGGGCATCCAGACCCCCCTCTCACCCCTCTGCT 1960  
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
DB 1961 CATGGTTCTCCAGCTCCTGTCTCTACTATGCTAAACGATGTCTTTCAGAGGGGCTCCA 2020  
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
DB 2021 AGGATGTCCCCAAGGCGCAGCATCTCTGAAATCACAGAGTTTCTCTCGGAGGGT 2080  
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
DB 2081 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGCTACTCCT 2140  
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680  
DB 2141 CCAGTAGCAAGGACGAGCTCCTCGGGGGAACGTGTCATCAGTGGTCAGTGGGGTTCCA 2200  
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
DB 2201 AGATTATCCCCTAAACACTATAGACCCAGGTCTCCAGACAGAACAGTATTGGAAATACC 2260  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
DB 2261 CCCAGTGGCCAGTTCTTGTCTTCTCCCAAGCTGTTATTTCCAACTGAACTGTGGC 2320  
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
DB 2321 ATGCCTATTCCAGCTGCATCTCTACGCTGTAGTCTCTGCATCGAACAGAGCTGTATCC 2380  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
DB 2381 CCTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCATCAGAGCAGAACTCTCTCGCAGG 2440  
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
DB 2441 AATAAGAAATAATTTAAACCCCAATGAACATCACCTAGTCTTCTCAAAAGCTCAAAACAAA 2500  
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
DB 2501 GGTATATCACCAAGTTTCTGAAACATAGAAACAGATTGATGATTAAAGAAATTTAAG 2560  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
DB 2561 AATGATTTTAGGTTACAGCAAGTTCTACTTCTGAAATCTATGGATCAACTACTATAACAAA 2620  
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
DB 2621 AATAGAGGGAGAAAAATCAAGAGATTTCATAAGACAAAAATTCAAACAGTGTCTAAG 2680  
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
DB 2681 GATTCCTTCATTTGAAATAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGC 2740  
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
DB 2741 CCCAGCATTTCCCTTCAATACTTAGTAACACGGAGCACAAGAGGGGACCTGAGGTCACT 2800  
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900  
DB 2801 TCCCAAGGGGTTACAGATTTCCAGCCAGCATGTAAACAAGAGAAACCATTAAGGAAGAG 2860  
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
DB 2861 AAGAAAGACGAGCTGAGCAAGTTAGGAAATCAACATTCGAATCCCAATGCAAGGAGTTTC 2920  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrProThrSerProArgProGln 940

2921	Db	A	ACCACAGTTCCTTCTCTCAGCCAAAGCGTTCTACTACCCCAACTTCACCTCGGCGCTCAA	2980
941	Qy	A	laGlnProSerProSerMetValGlnGlnProThrProValTyrThrGlnPro	960
2981	Db	G	CAACAACCTAGCCCATCTATGGTGGTCAACAACAGCCAACTCCAGTTTATACTCAGCCT	3040
961	Qy	V	alCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980
3041	Db	G	TTTGTGTTTGACCAAAATATGATGTATCCAGTCCAGTCGAGCCAGCGGTGCAACCTTTA	3100
981	Qy	T	yrProIleProMetThrProMetProValAsnGlnAlaIysThrTyrAlaValPro	1000
3101	Db	T	ACCCAAATACCTATGAGCCCATGCCAGTGAATCAAGCCAAGACATATAGACAGATACCA	3160
1001	Qy	A	snMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer	1020
3161	Db	A	ATATGCCCAACACGGGCACAGACGAGCATCATCAGAGTGCCATGATGACCCAGGTCA	3220
1021	Qy	A	laAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr	1040
3221	Db	G	CAGCGGGCCACCGATTGACGCCACCCACACAGCTTACTCCACGCAATATGTTGCCTAC	3280
1041	Qy	S	erProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln	1060
3281	Db	A	GTCTCTCAGCAGTTGCCAAATCAGCCCTGTTTTCAGCATGTGCCACATTTATCAGTCTCAG	3340
1061	Qy	H	isProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr	1080
3341	Db	C	ATCTCTAATGCTATAGTCTCTGTAATACAGGGTAACTGTAAGATGATGGCACCAACCA	3400
1081	Qy	H	isAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr	1100
3401	Db	C	ACGCCACCGCTGGTTTATGATCTCTTCAGCAACTCAGTACGGGGCTCATGACGACAGC	3460
1101	Qy	H	isAlaMetTyrAlaCysProGlyLeuProTyrAsnIysGluThrSerProSerPheTyr	1120
3461	Db	C	ATGCGATGATGATGATGTCCTCAAAATATACATACAAACAGGAGACAAGCCCTCTCTTCTAC	3520
1121	Qy	P	heAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis	1140
3521	Db	T	TTTGCCATTTCCAGGGGCTCCTTGCTCAGCAGTATGGCACCTTAAGCTACCTCGCTGCAC	3580
1141	Qy	P	roHisThrProHisProGlnProSerSerAlaThrProThrGlyGlnGlnGlnSerGlnHis	1160
3581	Db	C	CACATATCCACACCCCTCAGCTTACCTACCCCTCAGTACGACAGCAGCAAGCCAAACAT	3640
1161	Qy	G	lyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla	1180
3641	Db	G	GTGGAAGTCACTCTGCAACCCAGTCTGTTCAGCAGCACCATCAGCACCGAGCGCCAGGCT	3700
1181	Qy	L	euHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr	1200
3701	Db	C	TCCATCTGGCGAGTCCACAGCAGCAGTACGCCAATTTACACGCGGGGCTGCGCCAACT	3760
1201	Qy	P	roProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla	1220
3761	Db	C	CAACCTCCATGACACCTGCTCCACACGCGAGTGGCCACAGATAGTTTCCAGCAGACA	3820
1221	Qy	G	lnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro	1240
3821	Db	C	AAACAGACTGTCTTTACGATCCATCTCTCTCAGCTTACGGCGGTATACCAACCCACCC	3880
1241	Qy	H	isMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr	1260
3881	Db	C	ACATGGGCCACGATACCTTCAGGCTCATGTACAGTACAGATGGTTCCTCTCATTCCAACT	3940
1261	Qy	A	laHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu	1280
3941	Db	G	CCCATGCGCCAAATGATGCTAATGACGACACAGCCACCCGGCGGTCCCGAGCGCCCTC	4000
1281	Qy	A	laGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr	1300
4001	Db	G	CTCAAAATGACATACGCCCATTTCCAGTCTCGCAACACAGCGGCAATTTCCCTATATACG	4060

**Qy** 1301 HisProSerValGlnIalaHisHisGlnGlnGlnLeu 1312  
|||||  
**Db** 4061 CACCCTTCAGTACAAGCCCAACCAACAGCAGTTG 4096  
|||||

**RESULT 8**

HSDANSCA2	HSDANSCA2	4163 bp	mRNA	linear	PRI 18-APR-2005
LOCUS					
DEFINITION	H.sapiens mRNA for SCA2 protein.				
ACCESSION	Y08262				
VERSION	Y08262.1 GI:1770389				
KEYWORDS	SCA2 gene.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.					

**REFERENCE**

**AUTHORS** Imbert, G., Saudou, F., Yvert, G., Devys, D., Trottier, Y., Garnier, J.M., Weber, C., Mandel, J.L., Cancel, G., Abbas, N., Durr, A., Didierjean, O., Stevanin, G., Agid, Y. and Brice, A.

**TITLE** Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats

**JOURNAL** Nat. Genet. 14 (3), 285-291 (1996)

**PUBMED** 8896557

**AUTHORS** Imbert, G.

**TITLE** Direct Submission

**JOURNAL** Submitted (20-SEP-1996) G. Imbert, I.G.B.M.C., Departement Of Genetics, B.P. 163, 67404 Illkirch Cedex, FRANCE

**FEATURES**

source Location/Qualifiers

1..4163 /organism="Homo sapiens"  
/mol\_type="mRNA"  
/isolate="DAN patient"  
/db\_xref="taxon:9606"  
/cell\_line="lymphoblastoid"  
/clone\_lib="DAN"  
/dev\_stage="adult"  
<1..4163  
/genes="SCA2"  
<1..2747  
/genes="SCA2"  
/codon\_start=3  
/protein\_id="CAA69589.1"  
/db\_xref="GI:1770390"  
/db\_xref="GOA:Q99700"  
/db\_xref="UniProt/Swiss-Prot:Q99700"  
/translations="GNGGGAATPGSRRLGLGGPRFPVVLLPLASPGAPAAPTRA  
SPICARASPPRSVSLARPAAGCPRPACBPVVGPLTMSLKPOQQOQQOQQOQQOQQOQQO  
LQQRPNNSKGLPOSTISFGDIYANMRMHILTSVSGSKCEVOYKNKGIVGVFKTPGS  
CKDLVIDAAHEKSTESSSSGPKEIEIMESILFKCSDFVWVFQKMDPSYAKRAFTDSIA  
ISKAVNGEHKKOLEWDAGELTANLEALENDVSNQNDPMFYNSENYGVSVYTA  
DSLSYSYTVPLERDENSEEFLEKRAANLAEEIESAOYKARVALENDDRSREEKYTA  
VQRNRSREGHSLNTRENKYIPGQRNREVISWGSGRQNSPMMQPGSGSSMPRSRTSH  
TSDFNPNNSGDQRVNGVWPSPCPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP  
SFRESPPHPSAHSAPVSTPMKPMSSEGPPRMSPKAQRHPRNRHVSAGRGISGL  
EFVSHNPPEBAATPVVARTSPSGCTWSVGVPRLSPTKHPRSPRQNS IGNTPGSP  
VLASPOAGIIPTEAVAMPIPAASPTPASNAKRAVTPSAEKDSRLQDQRONSPAGNK  
ENTIKNETSPFSKAENKGISPVVSEHRKQIIDLKKFKNDRLQPSTSESMDQLNK  
NRGEKSRDLIKDKIPSADKSFIENSNNCTSGSSKPNPSLPSILNTEHKRGP  
VTSQGVOTSPACKQDDKEEKDAEQVRKSLTNPAKEKRNFRSPSQKPSPTTPTS  
PRPOAFSPSMWHQHQPTPVYTQPCFAPNMMYFVPVSPGVQYOICPNSGKTISIIRVP"

**gene**

**CDS**

1..348-88 Length: 4163  
6377..00 Matches: 1211  
Percent Similarity: 98.30% Conservative: 0  
Best Local Similarity: 98.30% Mismatches: 1  
Query Match: 91.61% Indels: 21

**ORIGIN**

Alignment Scores:  
Pred. No.: 1.34e-88  
Score: 6377.00  
Percent Similarity: 98.30%  
Best Local Similarity: 98.30%  
Query Match: 91.61%

DB:	8	1	Gaps:
US-10-802-228-2 (1-1312) x HSDANSCA2 (1-4163)			
QY	81	GlyAsnGlyGlyValAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyPro	100
DB	3	GGCAAGCGGGCGGCGGTTTCGGCCCGCTCCCGCGGCTCTCTGGTCTCGCGCGGCT	62
QY	101	ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProAlaAla	120
DB	63	CCCGGCCCTTCGTGCTGCTCTCTCCCTCCCGCGGCGGCGGCTCCCGCGGCG	122
QY	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu	140
DB	123	CCAACCGCGCTCCCGCTCGGCGGCTCCCGCGGCGGCGGCTCCCGCGGCTCTCTG	182
QY	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrglyProLeuThr	160
DB	183	GGCGGCCCGCTCCCGCTCGGCGGCTCCCGCGGCGGCGGCTCCCGCGGCTCTCTG	242
QY	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180
DB	243	ATGTCGCTGAGGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	302
QY	181	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
DB	303	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	362
QY	201	SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerSerSer	220
DB	363	AGCGGCTTCCTAGCGTCGCGCGCGCGCGCGCTTCGCGGCTTCGCGCTCTCTCG	422
QY	221	SerSerAlaThrAlaProSerValValAlaAlaThrSerGlyGlyGlyArgProGly	240
DB	423	TCTCGGCGCAGGCTCCTCTCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	482
QY	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260
DB	483	CTGGCAGAGGTGCAACAGTAACAAGAGTCCCTCAGTCTACGATTTCTTTTGATGA	542
QY	261	IleTyAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280
DB	543	ATCTATGCAATATGAGGATATATGAGGATATATGAGGATATATGAGGATATATG	602
QY	281	ValGlnValLysAsnGlyIleTyrglyValPheLysThrTyrglySerProLysCys	300
DB	603	GTCAAGTGAATAATGAGGATATATGAGGATATATGAGGATATATGAGGATATATG	662
QY	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg	320
DB	663	GAITTGCTACTTGATGCGCACATGAGAAAGTACAGATCCAGTTCGGGCGCGAAACGT	722
QY	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys	340
DB	723	GAAGAAATAATGAGGATATATGAGGATATATGAGGATATATGAGGATATATG	782
QY	341	AspMetAspSerTyrglyAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
DB	783	GATATGACCTCCAGTATGAGGATATATGAGGATATATGAGGATATATGAGGATAT	842
QY	361	ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla	380
DB	843	GTGAATGCGCAACACAAAGAGAGGACCTGGAGCCCTGGGATGCAGGTGAACCTCAC	902
QY	381	AsnGluIleLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMet	400
DB	903	AATGAGGAATTCAGGCTTTGGGAAATGACGTATCTAATGGATGGGATCCCAATGATG	962
QY	401	PheArgTyrglyAsnGluAsnTyrglyValValSerThrTyrglySerSerSerSer	420
DB	963	TTTCGATATATGAAGAAATATATGGGTATGTATGATGATGATGATGATGATGATG	1022
QY	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440

DB	1023	TATACAGTGCCCTTAGAAAGAGATAACTCAGAAGAAATTTTAAACGGAGCAAGGCA	1082
QY	441	AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrglyLysAlaArgValAlaLeuGlu	460
DB	1083	AACCAAGTTAGCAGAGAATAATTGAGTCAAGTGCCAGTACAAAGCTCGAGTGCCCTG	1142
QY	461	AsnAspArgSerGluGluGluLysTyrglyAlaValGlnArgAsnSerSerGluArg	480
DB	1143	AATGATGATAGAGTGAGAGAGAAATAACACAGCAGTTCCAGAGAAATTCAGAGTGA	1202
QY	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrglyProGlyGlnArgAsnArg	500
DB	1203	GAGGGGCAGCATAAACACATAGGAAATAATAATATTCCTCTCGGACAAAGAAATAG	1262
QY	501	GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520
DB	1263	GAAGTCATATCTGGGGAAGTGGAGACAGAAATTCACCGCGTATGGGCCAGCTGGATCG	1322
QY	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540
DB	1323	GGCTCATGCTCAAGATCCACTTCTCAGATTTCAGCCCGAAATTCCTGTTCA	1382
QY	541	AspGlnArgValValAsnGlyValProTrpProSerProCysProSerProSerSer	560
DB	1383	GACCAAGAGTAGTTAATGGAGGTGTTCCTGGCCATCGCTTGGCCATCTCTCTCTCT	1442
QY	561	ArgProProSerArgTyrglyGlnSerGlyProAsnSerLeuProArgAlaAlaThrPro	580
DB	1443	CGCCCACTTCTCGTACCAGTCAGGTCCCACTCTCTTCCACCTCGGCCAGCCACCT	1502
QY	581	ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla	600
DB	1503	ACACGGCGCCCTCCAGGCGCCCTCGCGCCATCCAGACCCCGCTCTCACCTCTCTCT	1562
QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
DB	1563	CATGGTCTCCAGCTCCTGTCTACTATGCTTAAACGATGCTTTCAGAGGCGCTCCA	1622
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
DB	1623	AGATGTCCTCCAAAGGCCAGGACATCTCCAAATCACAGAGTTTCTCTGGAGGGGT	1682
QY	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro	660
DB	1683	TCCATATCCAGTGGCTAGAAATTTGATCCCAACCCAGTCAAGCAGCTACTCTCT	1742
QY	661	ProValAlaArgThrSerProSerGlyThrTrpSerSerValValSerGlyValPro	680
DB	1743	CCAGTAGCAAGGACCAAGTCCCTCGGGGGAACGTGGTCTCATGTCAGTGGGGTTCCA	1802
QY	681	ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr	700
DB	1803	AGATTTCCCTTAAACTCATGACCCAGGTCTCCAGACAGACAGTATTGGAATACC	1862
QY	701	ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla	720
DB	1863	CCAGTGGGCGGCTCTCTCTCTCCCAAGCTGGTATTATTCCAACCTGAAGCTGTGCC	1922
QY	721	MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr	740
DB	1923	ATGCTTATTCAGCTCATCTCTACGCTGTGTAGTCTCTGATCGAACAGAGCTGTACC	1982
QY	741	ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly	760
DB	1983	CCTTCTAGTAGGCTTAAGATTCAGGCTTCAAGATCAGAGGAGCAACTCTCTCTGAGG	2042
QY	761	AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys	780
DB	2043	AATAAGAAATATTAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAA	2102
QY	781	GlyIleSerProValValSerGluHisArgLysGlnIleAspLeuLysLysPheLys	800

```
Db      2103 GGTATATCACCAGTTGTTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2162
QY      801  AnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db      2163 AATGATTTTAGGTTACAGCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAA 2222
QY      821  AnArgGluGlyGlyLysSerArgAspLeuLysAspLysLysIleGluProSerLalLys 840
Db      2223 AATAGAGGGGAGAAAATCAAGAGATTGTGATCAAAAGCAAAAATTTGAACCAAGTCTAAG 2282
QY      841  AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db      2293 GATTTCTTTTGAATAATAGACGACAACTGTACAGTGGCAGCAGCAAGCCGGAATAGC 2342
QY      861  ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db      2343 CCCAGCAVTTCCCTTCAATACTTAGTAACCGGAGCACAAGAGGGGACCTGAGGTCACT 2402
QY      881  SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGlyLysAspAspLysGluLys 900
Db      2403 TCCCAAGGGGTTTCAGACTTCCAGCCCGACATGTAAACAAGAGAAAGACGATAAGGAAGAG 2462
QY      901  LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db      2463 AAGAAAGCGCAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCCAAGAGATTC 2522
QY      921  AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db      2523 AACCCACAGTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCCCTCAA 2582
QY      941  AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db      2583 GCACAACTTAGCCCATCTATGGTGGGTCAATCAACAGCCCAACTCCAGCTTTATACTCAGCCT 2642
QY      961  ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db      2643 GTTTGTTTGGACCAATATGATGATTCAGTCCAGTCCAGGCGCCAGGGGTGCAATA---- 2698
QY      981  TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db      2699 -----CCCA 2701
QY      1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db      2702 AATATGCCCCAACACGCGCAGACGATCATCATCAGAGTGCCATGATCACCACCGCTCA 2761
QY      1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db      2762 GCAGCGGGCCCAACCGATTGCAGCCACCCACCAGCTTACTCCACGCAATATGTTGCCCTAC 2821
QY      1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db      2822 AGTCCTCAGCAGTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 2881
QY      1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaAlaArgMetMetAlaProProThr 1080
Db      2882 CATCCTCATGTCTATAGTCTGCTAATACAGGGTAATGTAGAAATGATGGCACCACCAACA 2941
QY      1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db      2942 CACGCCGACGCTGGTTTATGATCTCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3001
QY      1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db      3002 CATGCGATGATGATGATGCTCCCAAAATACCATACACAGGAGACAAAGCCCTCTTTCTAC 3061
QY      1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db      3062 TTTGTCATTTCCACGGGCTCCCTTCTCAGCAGTATGGCCACCTTAACGCTACCCCTGCAC 3121
QY      1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db      3122 CCACATATCCACACCTTCAGCTACCTACCCCTCAGTACCCCTCAGCAGCAGCAAGCAACAT 3181
```

```
QY      1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Db      3182 GGTGGAAGTCATCTCTGCACCAGTCTCTGTTTTCAGCACCATCAGCACCAGCCGCCAGGCT 3241
QY      1181 LeuHisLeuAlaSerProGlnGlnSerLysLysLysLysLysLysLysLysLysLys 1200
Db      3242 CTCCATCTGGCCAGTCCACAGCAGCAGTACGCCATTTACCAACGCGGGCTTTCGCCAAT 3301
QY      1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db      3302 CCACCTCCATCAGACCTGCTCCACACGAGCTGCCACAGAAATAGTTTCCCGACGCA 3361
QY      1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db      3362 CAACAGACTGTCTTTAGCATCATCTCTTCTCACGTTACGCGCGGTATACCAACCCACCC 3421
QY      1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db      3422 CACATGGCCCCAGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAAT 3481
QY      1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db      3482 GCCCATGGCCCAATGATGCTAATGACGACACACGACCCCGCGGTCCCGAGCGCCCTC 3541
QY      1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db      3542 GCTCAAAAGTGCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTTATATGACG 3601
QY      1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
Db      3602 CACCTTTTCAGTACAAAGCCCAACCAACAGCAGTTG 3637

RESULT 9
LOCUS   A62706                4200 bp          DNA          linear          PAT 12-MAR-1998
DEFINITION   Sequence 7 from Patent WO9717445.
ACCESSION   A62706
VERSION     A62706.1  GI:3716590
KEYWORDS    .
SOURCE      .
ORGANISM    unidentified
            unclassified sequences.
REFERENCE   1
  AUTHORS   Tora,L., Lutz,Y., Trottier,Y., Mandel and Jean-Louis.
  TITLE     METHOD FOR TREATING NEURODEGENERATIVE DISEASES USING A 1C2 ANTIBODY
            OR A FRAGMENT OR DERIVATIVE THEREOF, AND CORRESPONDING
            PHARMACEUTICAL COMPOSITIONS
  JOURNAL   Patent: WO 9717445-A 7 15-MAY-1997;
            CENTRE NAT RECH SCIEN (FR)
            Other publication FR 2741088 19970516.
FEATURES    Location/Qualifiers
            source
              1..4200
              /organism="unidentified"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32644"
              /clone="DAN1"

ORIGIN

Alignment Scores:
Pred. No.:      1.35e-88      Length:      4200
Score:          6377.00      Matches:    1211
Percent Similarity: 98.30%      Conservative: 0
Best Local Similarity: 98.30%      Mismatches: 1
Query Match:    91.61%      Indels:    21
DB:              6              Gaps:      1

US-10-802-228-2 (1-1312) x A62706 (1-4200)

QY      81  GlyAsnGlyGlyGlyAlaPheArgProGlySerArgLeuLeuGlyLeuGlyPro 100
Db      3  GGCAACGGCGCGCGCGTTCGGCGCCGCTCCCGCGGCTCTCTTGGTCTCGCGGGGCT 62
```



Db 2223 AATAGAGGGGAGAAAAATCAAGAGATTGTGATCAAAAGACAAAATTGAACCAAGTCTTAAG 2282

Qy 841 AspSerPheIleGluAsnSerSerSerAenCysThrSerGlySerSerLysProAsnSer 860

Db 2283 GATTCTTTTCATTTGAAAATAGCAGCAGCAACTGTACCAAGTGGCAGCAGCAAGCCGAATAGC 2342

Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880

Db 2343 CCCAGCATTTCCCTTCAATACTTAGTAACCGAGACCAAGAGGGGACCTGAGGTCACT 2402

Qy 881 SerGlnGlyValGlnThrSerSerProLaCysLysGlnGluLysAspAspLysGluGlu 900

Db 2403 TCCCAAGGGTTTCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGATAAGAGAGAG 2462

Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920

Db 2463 AAGAAAGCGCAGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAGAGGTTTC 2522

Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940

Db 2523 AACCCACGTTCTTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCTCAA 2582

Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960

Db 2583 GCACAACTTAGCCCATCTATGTGTGGTCAATCAAGCCCACTCCAGTTTATCTCAGCCT 2642

Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980

Db 2643 GTTTGTTTGGCACCATAATATGATGTATCCAGTCCAGTGAGCCAGCGGTGCAATA --- 2698

Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000

Db 2699 -----CCAA 2701

Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020

Db 2702 AATATGCCCAACAGCGGACAGCAGCATCATCAGAGTGCCATGATGACCACCGGTCA 2761

Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040

Db 2762 GCAGCGGGCCACCGATTGTCAGCCACCCACCGACTTACTCCACGCAATATGTTGCCTAC 2821

Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060

Db 2822 AGTCCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 2881

Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaAlaMetMetAlaProProThr 1080

Db 2882 CATCTCATGCTATAGTCTGTATACAGGTAATGCTAGAAATGATGCACCAACCA 2941

Qy 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100

Db 2942 CACGCCCGCCTGTTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGAGC 3001

Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120

Db 3002 CATGCGATGTATGCATGTCCCAAAATACCATACAAGAGAGCAAGCCCTTCTTTCTAC 3061

Qy 1121 PheAlaIleSerThrClySerLeuAlaGlnTyrAlaHisProAsnAlaThrLeuHis 1140

Db 3062 TTTGGCATTTCCAGCGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTCGAC 3121

Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160

Db 3122 CCACATACTCCACACCTCAGCTTCAGTACCCCACTGGACAGCAGCAAGCCAAAT 3181

Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180

Db 3182 GGTGGAAGTTCATCTCGACCCAGTCTGTTCAGCACCATCAGCACAGCCGCCAGGCT 3241

Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200

Db 3242 CTCCTATCTGGCAGTCCACAGCAGCAGTCAAGCATTTTACCAGCGGGGCTTCGGCCAAC 3301

Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220

Db 3302 CCACCTCCATGACACCTGCTCCCAACACGAGTCCGACAGAATAGTTTCCAGCAGCA 3361

Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240

Db 3362 CAACAGACTGTCTTTAGCATCCCTTCTCAGCTTACGCGGTATACCAACCCACCC 3421

Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260

Db 3422 CACATGCCCACTGACCTCAGGCTCATGACAGTCAAGTGGTTCCTTCTCATCCAAC 3481

Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280

Db 3482 GCCCATGCCCAATGATCTAATGACGACACAGCACGCCGCGGTCCCGCCGCTC 3541

Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300

Db 3542 GCTCAAAAGTGACACTACAGCCCATTCAGTCTCGAACACAGCGCATTTCCCTATATGACG 3601

Qy 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312

Db 3602 CACCTTCAGTACAAGCCCAACCAACAGCAGTTG 3637

RESULT 10

AF041472

LOCUS

DEFINITION Mus musculus ataxin-2 (SCA2) mRNA, linear ROD 28-NOV-2001

ACCESSION AF041472

VERSION AF041472.1 GI:3005019

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 4225)

AUTHORS Nechiporuk, T., Huynh, D.P., Figueroa, K., Sabha, S., Nechiporuk, A. and Pulst, S.M.

TITLE The mouse SCA2 gene: cDNA sequence, alternative splicing and protein expression

JOURNAL Hum. Mol. Genet. 7 (8), 1301-1309 (1998)

PUBMED 9668173

REFERENCE 2 (bases 1 to 4225)

AUTHORS Nechiporuk, T.T., Figueroa, K., Sabha, S., Nechiporuk, A.V. and Pulst, S.M.

TITLE Direct Submission

JOURNAL Submitted (07-JAN-1998) Medicine/Neurology, Cedars-Sinai Medical Center, 8700 Beverly Blvd., Los Angeles, CA 90048, USA

FEATURES

Location/Qualifiers

1..4225

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/chromosome="12"

/map="12q23.1"

1..4225

/gene="SCA2"

27..3884

/gene="SCA2"

/codon\_start=1

/product="ataxin-2"

/protein\_id="AAC09275.1"

/db\_xref="GI:3005020"

/translation="MRSSTAAVQRPAAQDPEPRPAGWAARRSLPRTARRGGRGAVAYPSAGPPRGPGAPGRPSPPSCDCFSNGHSGASRPSRRLLGVCGPPRPFFVVVLLALAPAATPRACPPGVRASPPRSVSSARPAPCPACPEPVGPTLMSLKPPQPPAPATGRKPGGGLSSPGAAPASAVTSASVPAAPAVASSAAAGGSRGGLGRGNS SKGLPQPTISPDGIYANVRMVHLTSSVGSKEVQKNGIYEGVFTKYSKCDLVD AHAESKTESSSGPKREIMESVLFKCSDFVVQFKDTSVYARDAFTDSALSAKNG EHKKDLPTDAGSLTASELELNDVSGNDPNDMPRYNEENGVVSTYSSLSSTY VPLERDINSBEFLKREARNAQLAEIESSAQYKARVALENDDRSEEEKYTAVORNSDR

gene

CDS





QY 596 SerHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMetSer 615  
DB 1722 TCTCACCCCTCTGCTCATGGTTCTCCAGCTCTGCTCTACTATGCCCTAAACGCGATGCT 1781  
QY 616 SerGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgVal 635  
DB 1782 TCAGAAGGACCCCAAGGATGCTCCAAAGGCACAGCGCCACCCTCGGAATCACAGATC 1841  
QY 636 SerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSer 655  
DB 1842 TCTGCTGGGAGAGGCTCCATGCTCTAGTGGCTAGAAATTTGATCCCAAAATCCCAAGT 1901  
QY 656 GluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerSerVal 675  
DB 1902 GAAGCAGCTGCTCTCCAGTGGCAAGGACCAAGTCTCGCAGGGGGAAACGTGGTCTCTCAGTG 1961  
QY 676 ValSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsn 695  
DB 1962 GTCAGTGGGGTTCCAAAGGTATCTCCCAAACTCACAGCCAGGCTCTCCAGGCAGAGC 2021  
QY 696 SerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleIlePro 715  
DB 2022 AGCATTTGNAACCTCTCCAGCGGGCTGTGCTTGTCTTCTCCCAAGCTGGCATCATCCCT 2081  
QY 716 ThrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSer 735  
DB 2082 GCAGAAGCGGTTTCCATGCTCTGCCCGCGCATCTCCGACTCTCGCCAGCGCTGCATCC 2141  
QY 736 AsnArgAlaValThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGln 755  
DB 2142 AACAGAGCACTGACCCCATCTATTGAGGCACAAAGATTCCAGCTTCAAGATCAGAGGCAG 2201  
QY 756 AsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSer 775  
DB 2202 AACTCTCTGCGAGGAGTAAGAAATGTTAAGCAAGTGAACATCATCCTAGCTTTTCA 2261  
QY 776 LysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAsp 795  
DB 2262 AAAGCTGCACAAAGGTATGTCACAGTTGTTCTGAAACACAGAAAAACAGATTGATGAC 2321  
QY 796 LeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAsp 815  
DB 2322 TTAAGAAAGATTTAAGAAATTTAGGTATACGCCAAGCTTACATCTGAAATCTATGGAT 2381  
QY 816 GlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIle 835  
DB 2382 CNACTACTAAGCAAAAATAGAGAGGAGAAAAGTCCAGAGATTGATTAAAGATAAACG 2441  
QY 836 GluProSerAlaLysAspSerPheIleGlu-----AsnSerSerSerAsnCysThrSer 853  
DB 2442 GAAGCAAGTGCTAAGGATAGTTTTCATTGACAGCAGCAGCAGCAGCAACTGTATCCAGT 2501  
QY 854 GlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThrGluHis 873  
DB 2502 GGCAGCAGCAACAAACAGCCCTAGCATCTCCCTTCCATGCTTAGTAATGCAGAGCAC 2561  
QY 874 LysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCysLysGln 893  
DB 2562 AAGAGGGGGCTGAGGTACATCCCAAGGGGTGCAGACTTCCAGCCCAAGCTTCGCAACAA 2621  
QY 894 GluLysAspAspLysGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeu 913  
DB 2622 GAGAAAGGATGACAGAGAGAAGAAGACACAACACAGAGCAGGTTAGGAAATCGACATTG 2681  
QY 914 AsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLysProSerThrThr 933  
DB 2682 AATCCCAATGCAAGAGGATTCACCCCTCGTTCTTCTCTCAGCCCAAGCCCTTCTACTACC 2741  
QY 934 ProThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnPro 953  
DB 2742 CCNAGCTCACCTCGGCCTCAAGCACAAACCAGCCCATCTATGGTGGGTTCATCAGAGCCA 2801

QY 954 ThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal 973  
DB 2802 GCTCCAGGTACACTCAGCCTGTGTGCTCGCACCAATATGATGTATCCCGTCCCAAGT 2861  
QY 974 SerProGlyValGlnProLeuTyrProIleProMetThrProMetProValAsnGlnAla 993  
DB 2862 AGCCCGGGCGTACAAACCTTTATACCCAAATACCTATGACGCCATGCTGTGAAACAAGCC 2921  
QY 994 LysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAspGlnHisHis 1011  
DB 2922 AAGACATATAGACGAGGTAAAGTACCAAAATATGCCCCAACAGCAGCAAGACCAACATCAT 2981  
QY 1012 GlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThrProPro 1031  
DB 2982 CAAAGCACCATGATGCACCCAGCCTCGCGCAGGGCCACCCTCGTAGCCACCCCGGCC 3041  
QY 1032 AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGlnProLeuVal 1051  
DB 3042 GCTTACTCCACTCAGTACGTTGCTCAGGCCCTCAGCAGTTTCCCAATCAGCCTTTGGTC 3101  
QY 1052 GlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGly 1071  
DB 3102 CAGCATGTGCGCATTTATCAGTCTCAGCATCCTCATGTGTACAGTCTCTGTATACAAGT 3161  
QY 1072 AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAla 1091  
DB 3162 AATGCCAGGATGATGGCACCACCAGCAGCATGCTCAGCCTGGTTAGTGTCTTCTTCAAGT 3221  
QY 1092 ThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyr 1111  
DB 3222 GCTCAGTTCGGGGCTCAGAGCAGCAGCCCATGTATGCTATGCTCCAAATTACCATAC 3281  
QY 1112 AsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGln 1131  
DB 3282 AACAAAGAGACAAAGCCCTCTTTCTACTTTGGCAATTTCCACCGCTCCCTCGCTCAGCAG 3341  
QY 1132 TyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThr 1151  
DB 3342 TATGCATATCTTAATGCGCGCTGCATCCACATATCCCCATCTCCAGCCTTCGGCCACT 3401  
QY 1152 ProThrGlyGlnGlnSerGlnHisGlyLysSerHisProAlaProSerProValGln 1171  
DB 3402 CCCACCGCAGCAGCAGAAAGCCAGCATGGTGAAGTACCCTGCACCCAGCTCTGTTCAG 3461  
QY 1172 HisHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGlnGlnSerAla 1191  
DB 3462 CACCATCAGCACCAAGGCTGCCAGGCTTCTCATCTGGCCAGTCCACAGCAGCAGTCGGCC 3521  
QY 1192 IleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSerAsnThrGln 1211  
DB 3522 ATTTATCATGCGGGCTGGCACCACCAACACCTTCCATGACACCTGCTCTTAATACACAG 3581  
QY 1212 SerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHis 1231  
DB 3582 TCTCCACAGACAGTGTTCACAGCAGCACAACAGCAGTCTTTCACCATCCACCTTCTCAT 3641  
QY 1232 ValGlnProAlaTyrThrAsnProHisMetAlaHisValProGlnAlaHisValGln 1251  
DB 3642 GTTCAGCGGGATACACCCACCCACCCACCATGCCCCAGTACCTCAGGCTCATGTACAG 3701  
QY 1252 SerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMetThrThrGln 1271  
DB 3702 TCAGGAATGGTTCCTTCTCATCCAACTGCCCATCGCCCAATGATGTCTAATGACGACAG 3761  
QY 1272 ProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProIleProValSer 1291  
DB 3762 CCACCC---GGTCCCAAGCGCCCTCGCTCAAAAGTAGTGACTACAGGCCATTTCCAGTTTCG 3818  
QY 1292 ThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaHisHisGlnGlnGln 1311  
DB 3819 ACAACAGGCATTTCCCTTATATGACGCAACCTTCAGTACAAAGCCCAACCAACAGCAG 3878  
QY 1312 Leu 1312



```
QY 705 ValLeuAlaSerProGlnAlaGlyllelleProThrGluAlaValAlaMetProIlePro 724
DB 1622 GTGCTTGCTTCTCCCAAGCTGGCATCTCCCTGCAGAAAGCGTTTCCATGCTGTGCC 1681
QY 725 AlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGlu 744
DB 1682 GCCGATCTCCGACTCTCCGAGCCCTGCATCCACAGAGCACTGACCCCATCTATTGAG 1741
QY 745 AlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsn 764
DB 1742 GCAAAAGATTCCAGGCTTCAAGATCAGAGGCAGACTCTCTCGCAGGGAGTAAGAAAT 1801
QY 765 IleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerPro 784
DB 1802 GTTAAAGCAAGTGAACATCACTAGCTTTTCAAAAGCTGCACAACAAAAGTATGTCA 1861
QY 785 ValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArg 804
DB 1862 GTTGTTCTGAAACACAGAAACAGATTGATGACTTAAAGAGTTTAAAGATTATTAGG 1921
QY 805 LeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGly 824
DB 1922 TTACAGCCAAGCTCTACATCTGAATCTATGGATCAACTTAAGCAAAATAGAGAAGA 1981
QY 825 GluLysSerArgAspLeuIleLysAspLyslleGluProSerAlaLysAspSerPheIle 844
DB 1982 GAAAAGTCACAGAGATTGATTAAAGATAAAACGGAAGCAAGTGTCTAAGGATAGTTTCAT 2041
QY 845 Glu-----AsnSerSerSerAsnCythrSerGlySerSerLysProAsnSerProSer 862
DB 2042 GACACAGCAGCAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCAACAGCCCTAGC 2101
QY 863 IleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGln 882
DB 2102 ATCTCCCTTCCATCTTAGTAATCGACAGCACAGAGGGGCTGAGGTACATCCCAA 2161
QY 883 GlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluLysLys 902
DB 2162 GGGGTGCAGACTTCCAGCCAGCTTCGCAAAACGGAAGAGGATGACAGAGAAGAGAAA 2221
QY 903 AspAlaAlaGluValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnPro 922
DB 2222 GACACAAACAGACAGCGTTAGGAAATTCGACATTGAAATCCCAATGCAAGAGGTTCAAC 2281
QY 923 ArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGlnAlaGln 942
DB 2282 CGTTCTTCTCAGCCCAAGGCTTCTACTACCCCAACGTCACCTCGGCTCAAGCACA 2341
QY 943 ProSerProSerMetValGlyHisGlnGlnProThrProValTyThrGlnProValCys 962
DB 2342 CCCAGCCCATCTATGGTGGGTCAATCAGCAGCCAGCTCCAGTGTACACTCAGCCTGTGT 2401
QY 963 PheAlaProAsnMetMetTyProValProValSerProGlyValGlnProLeuTyPro 982
DB 2402 TTCGCACCCCAATATGATGATCCCGTCCAGTGAGCCGGGGGTACAACCTTTATACCCA 2461
QY 983 IleProMetThrProMetProValAsnGlnAlaLysThrTyArgAla-----ValPro 1000
DB 2462 ATACCTATGACGCCCATCGCTGTGNAACCAAGCCAGACATATAGAGCAGGTAAGTACCA 2521
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
DB 2522 AATATGCCCAACAGCGCAAGACCAACATCATCAAGCACCATTGATGATGATGATGATGAT 2581
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyThrSerThrGlnTyThrValAla 1040
DB 2582 GCGGAGGGGCCACCATCGTAGCCACCCCGCGCTTACTCCACTCAGTACGTTGGCTTAC 2641
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyThrGlnSerGln 1060
DB 2642 AGCCCTCAGCAGTTTCCCAATCAGCCTTGTGTCAGCATGTGCCCATATCAGTCTCAG 2701
```

```
QY 1061 HisProHisValTySerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 2702 CATCTCATGTGTAGTACGTCTGTATACAAGGTAAATCCAGGATGATGGCACCCACGCA 2761
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyThrGlyAlaHisGluGlnThr 1100
DB 2762 CATCTCAGCTGGTGTAGTGTCTTTCAGCTGCTCAGTTCGGGGCTCAGAGCAGACG 2821
QY 1101 HisAlaMetTyAlaCysProLysLeuProTyThrAsnLysGluThrSerProSerPheTy 1120
DB 2822 CACGCCATGTATGATGATGTCCCAATATACCATACAAGAGAGACAAGCCCTTCTTCTAC 2881
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyThrAlaHisProAsnAlaThrLeuHis 1140
DB 2882 TTTGCCATTTTCCACGGCTCCTCGCTCAGCAGTATGCACATCCTAATATGCGCCCTGCA 2941
QY 1141 ProHisThrProHisProGlnProSerSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
DB 2942 CCACATACTCCCATCTCCTCAGCTTCGGCCACTCCACCGGACAGCAGCAAGCCAGCAT 3001
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180
DB 3002 GGTGGAAGTCACCTCGACCCAGTCTGTTTCAGCACCATCAGCACCAGGCTGCCAGGCT 3061
QY 1181 LeuHisIleAlaSerProGlnGlnSerAlaIleTyThrHisAlaGlyLeuAlaProThr 1200
DB 3062 CTTTCATCTGGCAGTCCACAGCAGCAGTGGGCCATTTATCATGCGGGGCTGGCACCACA 3121
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnSerProGlnSerPheProAlaAla 1220
DB 3122 CCACCTTCCATGACACCTGCTCTAATACACAGTCTCCACAGAGCAGTTCCTCCAGCAGCA 3181
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyThrAsnProPro 1240
DB 3182 CAACAGACAGTCTTCCACATCCACCTTCTCATGTTTCAGCGGATATACACCCACCCACC 3241
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3242 CACATGGCCCAAGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAAC 3301
QY 1261 AlaHisAlaProMetMetLeuMetThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
DB 3302 GCCCATGGCCCAATGATGCTAATGACACACAGCCACCC-----GGTCCCAAGGCCGCCCTC 3358
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyMetThr 1300
DB 3359 GCTCAAAGTGCACTACAGCCCATTCAGTTCGACACAGCGCATTTCCCTTATATGAGC 3418
QY 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
DB 3419 CACCTTTCAGTACAAAGCCCAACCAACAGCAGTTG 3454
```

## RESULT 12

```
LOCUS BC097692 2558 bp mRNA linear VRT 02-JUN 2005
DEFINITION Xenopus laevis cDNA clone MGC:115230 IMAGE:5079818, complete cds.
ACCESSION BC097692
VERSION BC097692.1 GI:66910767
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE
AUTHORS 1 (bases 1 to 2558)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE
AUTHORS 2 (bases 1 to 2558)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
```

Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

3 (bases 1 to 2558)

Klein, S. and Gerhard, D.S.

Direct Submission

Submitted (01-JUN-2005) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA

NIH-MGC Project

Contact: XGC help desk

Email: cgabps@mail.nih.gov

Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 229 Row: h Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

## FEATURES

## source

1. .2558

/organism="Xenopus laevis"

/mol\_type="mRNA"

/db\_xref="taxon:8355"

/clone="MGC:115230 IMAGE:5079818"

/tissue\_type="Oocytes"

/clone\_lib="NICHG XGC\_001"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6.1"

228..2120

/codon\_start=1

/product="Unknown (protein for MGC:115230)"

/protein\_id="AAH97692.1"

/db\_xref="GI:66910768"

/translation="MSMKQAGNNRKPGGSVPAAGGAGGGGRQNMGRGRIISQKGPQ

## CDS

SISFDGIYANRLVHLTSVSGKCEVLVNGNVYEGVKTYSKCDLVLDAHKTT

ESIVGPKRDIIVDFKSDVFMQVDFKMDVNVARRDFTDSGIACRLNGSHKEDL

EPWDGESLISPTQKLESLSENSGNDPMENRYENRYGVVSTYDSSIVTY

PLRDNSEELYKAEAAQAEIESSQVKARVALENDERSEEEKYTAQVALNDQE

RHSLNTRNKYIIPVORNDQAMLWTTHSGRONSARIGQSGSGPIPSRTGSHSDYNP

AGADQVNGGVWNPSPCPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP

SP

AGASVQVPAPVSTMPKRMSSGPPRMSPAQKHPKTHRVNMGRTMCGGLFVSNE

PPESVSMVPAPASPTPASDRAVTPSCAEKDTQLQDQRQNSPAGCRENSKYSSEK

TENKAVSPMASEQRKQLDDLLKKFKNDRLQSPSSPDALDQLPVKSRDSVEKPRDLKK

KNGFKC"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,78e-32 Length: 2558

Score: 2612.00 Matches: 527

Percent Similarity: 79.35% Conservative: 80

Best Local Similarity: 68.89% Mismatches: 85

Query Match: 37.52% Indels: 73

DB: 5 Gaps: 15

US-10-802-228-2 (1-1312) x BC097692 (1-2558)

QY 155 ValTyrGlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGln 174

213 GTATATGGGACCGTC---ATGTCCATGAAGCAGCAG----- 245

QY 175 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 194

246 -----CCCGGTAAAC 254

QY 195 ValArgLysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSer 214

255 AACAGAAAGCCCGCGGGGCGAGCGTC-----CCGGCTGCT----- 290

QY 215 SerSerSerValSerSerSerSerAlaThrAlaProSerSerValValAlaThrSer 234

291 -----GGAGGAGGAGCA 302

QY 235 GlyGlyGlyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSer 254

303 GGAGGAGGAGGAGCAGAAATATGGCGAGGAGGAGGATCAGTGGAAAGGGCCCTCAATCG 362

QY 255 ThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerVal 274

363 ---ATCTGTTTGGTGGTATCTATGCAAAATATGACACTGGTTCATATTTCTACATCAGTA 419

QY 275 ValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLys 294

420 GTTGCTCAAAATGTGGGTATTGGTGAAAAATGAAACCTATATGAAGGTGTTTTTAA 479

QY 295 ThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSer 314

480 ACTTACAGTCTTAAGTGGCATTTGGTGTGTTGATGAGCTCATAGAAAACTACAGAAATCT 539

QY 315 SerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPhe 334

540 ATTGTGGGCGCAAGCGGTGAAGACATTTGGGATAGTATATATTGTTTAAATCTTCCGACTTT 599

QY 335 ValValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAsp 354

600 GTTATGTTGCAGTTTAAAGATATGGATGTGAATTTATGCAAGACGAGATGGGTTTCACAGAT 659

QY 355 SerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrpAsp 374

660 TCTGGAATTTGCTTGTAGTTAAATGGTGAACATAAGGAGAGATCTGGAGCGCGTGGGAT 719

QY 375 AlaGlyGluLeu-----ThrAlaAsnGluGluLeuGluAlaLeuGlu 388

720 GTGTGTGAAGTCTCATTTAGCAGCACCTCTGCGAAAGATGAGCTGGAATCTTTGGAA 779

QY 389 AsnAspValSerAsnGlyTyrAspProAsnAspMetPheArgTyrAsnGluGluAsnTyr 408

780 ATGAG---TCTAATGATGGGATCCAATGAATGATTTTCGTACATCAAGAAAAATTAT 836

QY 409 GlyValValSerThrTyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAsp 428

837 GGTGTGTGTCAACTTATGACAGTAGTCTGTCTCTGTACACTGTTCCATTAGAAAGGAT 896

QY 429 AsnSerGluGluPheLeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluLeuGlu 448

897 AATTCAAGAAATATTGTAAGAGAGAGAGCCAGACGAGCTCAGATCTCTGAAGAAATTGAA 956

QY	449	SerSerAlaGlnTyrLysAlaArgValAlaLeuGluAAsnAspArgSerGluGlu	468
DB	957	TCAAGTTCCTCCAGTATAAAGCTAGAGTTGCTTGGAAATGATGCGGATCAGAGAGAA	1016
QY	469	LysTyrThrAlaValGlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArg	488
DB	1017	AAATACACGGCTGTTCACAGAGCCCTTAATGACCAGGAAGCCACTCACTGAATACCCAGA	1076
QY	489	GluAsnLysTyrIleProProGlyGlnArgAsnArgGluValIleSerTrp	506
DB	1077	GAATATAGTACATTCACCTGTCGAGAGAAATAAGATGCATGTTATGGACTACTCAC	1136
QY	507	SerGlyArgGlnAsnSerProArgMetGlyGlnProGlySerGlySerMetProSerArg	526
DB	1137	AGTGAAGGCAAAATCTGCAAGGATAGACAGCTCTGGATCAGGACCCATTCTTCAAGG	1196
QY	527	SerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnArgValValAsn	546
DB	1197	ACAGGATCACATTCCTCTGACTACAAACCCCAATGCTGGAGCTGATCAGAGAGTTGTAAAT	1256
QY	547	GlyGlyValProTrpProSerProCysProSerProSerSerArgProProSerArgTyr	566
DB	1257	GGAGTGTTCCCTGGCCATCGCTTGTCACTCTCTTCCCGCCACTTCTCGCTAC	1316
QY	567	GlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrArgProProSerArg	586
DB	1317	CAGTCAGGTCCCAACTCTCTCCACTCGGGCAGCACCCCTTACAAGGCGCCCTCAAGG	1376
QY	587	ProProSerArgProSerArgProProSerHisProSerAlaHisGlySerProAlaPro	606
DB	1377	CCCCCTTCGGGCCCATCAGGCCCGCTCTCACCTCTGCTCATGTTCTCCAGCTCCT	1436
QY	607	ValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerProLysAla	626
DB	1437	GTCCTACTATGCTTAACGCATGCTTCAGNAGGACTCCAAGATGTCCTCAAGCCA	1496
QY	627	GlnArgHisProArgAsnHisArgValSerLaglyArgGlySerIleSerSerGlyLeu	646
DB	1497	CAACGTCTACTCTCGTACTCACAGAGTTTCAAAATGCGAGGGGCAACAATGTGGAGGATTA	1556
QY	647	GluPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArgThrSer	666
DB	1557	GAATTTGTTCTCGCAATGAAGCTGGAGAGCGTCTGTGCACACTGTGGCCCGGAATAGT	1616
QY	667	ProSerGlyGlyThrTrpSerSerValValSerGlyValProArgLeuSerProLysThr	686
DB	1617	TCCTCAGTGGGAACATGCTCTCTGTTGAGCGGAGTTCAAGGTTATGCCCAAGACA	1676
QY	687	HisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyProValLeu	706
DB	1677	CACAGGCCAAGTCTCCGCGG-----CTAGGAAGTCTGACACCAAGTTCAAATGCG	1727
QY	707	AlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAla	726
DB	1728	ACTTCGCCCCAGCTGGGACAGTGCCTCCGAGTCTGTTTATATGCTGAGTCCCTGCCCA	1787
QY	727	SerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLys	746
DB	1788	TCCCTTACACCTGTAGCCCTGCATCTGATAGGCTGTGTACTCCCTCTGTGAGCAAAA	1847
QY	747	AspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnLys	766
DB	1848	GATACAAGACTACAAGATCAGACAGAAAAATCCCCAGCTGGTGTGTAAGAAAAATAGTAAA	1907
QY	767	ProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerProValVal	786
DB	1908	TACCGCAATCT-----AGCAAACTGAAAAACAAGCAGTATCTCCAATGGCT	1955
QY	787	SerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGln	806
DB	1956	TCTGAACAAAGAAACAGCTTGATGATTTAAAGAAATTTAAAAATGATTTTAGGCTACAG	2015
QY	807	ProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGly---	825
DB	2016	CCCAAGCTCAAGTCTCTGATGCTTTGGATCAGTTACCCGTCAAAAGTCGGGATAGTGAGAA	2075
QY	826	LysSerArgAspLeuIleLys-AspLysIleGluProSerAlaLysAspPheIleGln	845
DB	2076	AAGCAAGGGATCCGTTGAAGAAAAAATGGATTTCAAAATGCTAAAGAAATCAACAAGTGA	2135
QY	845	uAsnSerSerSerAsnCysThrSer---GlySerSerLysProAsnSerProSerIleSe	864
DB	2136	AAGCAGCAGTAATCCCACTAGCAACAACAGCAGCAAGCATGCAGTCCAGCATTTTC	2195
QY	864	rProSerIleLeu---SerAsnThrGluHisLysArgGlyProGluValThrSerGlnGln	883
DB	2196	ACCATCTATTATCAGCAGCAGTCTCTGAACAACAAGAGAGGTCAGAGGTAACATCACAAGG	2255
QY	883	yValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu---	902
DB	2256	TGTACAGACCTCTGGGCTTCC---AAACAGGATCGTGATGACAAGGATGACATGAAGGA	2312
QY	902	sAspAlaAlaGlu	906
DB	2313	AAATGCTGCAGAG	2325
RESULT 13			
AK095017			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1			
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,			
Wakamatsu, A., Hayaishi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,			
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,			
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,			
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,			
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,			
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,			
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,			
Tanikawa, M., Yanazaki, M., Ninomiya, K., Ishibashi, T., Yanashita, H.,			
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,			
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,			
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,			
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,			
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,			
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,			
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,			
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,			
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,			
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,			
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,			
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,			
Ozaki, K., Hirao, M., Omori, Y., Kawabata, A., Hiki, T., Kobatake, N.,			
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,			
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,			
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,			
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,			
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,			
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,			
Nakai, T., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.			
Complete sequencing and characterization of 21,243 full-length			
human cDNAs			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
1			
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,			

Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yanashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
3 (bases 1 to 2766)  
Isogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

## FEATURES

source

1. .2766  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BRHIP2015679"  
/tissue\_type="hippocampus"  
/clone\_lib="BRHIP2"  
/note="cloning vector: pME18SFL3"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.56e-28 Length: 2766  
Score: 2380.00 Matches: 497  
Percent Similarity: 71.23% Conservative: 13  
Best Local Similarity: 69.41% Mismatches: 69  
Query Match: 34.19% Indels: 138  
DB: 8 Gaps: 11

US-10-802-228-2 (1-1312) x AK095017 (1-2766)

QY 6 AlaAlaProArgSerProAlaValAlaThrGluSerArg 2766  
Db 19 TCAGCCCGGCTCCCTCCGACGATCTCCGCCGCGAGCGTGGGGATGCTCGGGCAG 78  
QY 19 -----ArgPheAlaAlaAArgTTP----- 25  
Db 79 CTCTCCACTCCGGTTTAGGTGTGAACGTTGAGGGGCTGAGGCTGTGGTGGCGTTT 138  
QY 26 -----ProGlyTTP-----ArgSerLeuGlnArgPro 34  
Db 139 TCGGGAACATGTCCTCCATCGGGGACATCTCTGGAGGGGAGAGTTAGGCCCGCTC 198  
QY 35 AlaArgArgSerGlyArgGlyGlyGlyAlaAlaPro-----GlyProTyrPro 51  
Db 199 CCCGTGCGGTAAAGGGGTAGGACCGCGCTCTCCGGAATCATCAGGCTGTGCGGG 258  
QY 52 SerAlaAlaProProProGlyProGlyProProProProSerArgGlnSerProPro 71  
Db 259 GCTCTCTCCCGCCCTCCGAGTCTCTGGGAAAGAT----- 294  
QY 72 SerAlaSerAspCysPheGlySerAsnGlyAlaPheArgProGlySer 91  
Db 295 -----CGAGGACGGGG 306  
QY 92 ArgArgLeuLeuGlyGlyGlyPro---ProArgProPheValValLeuLeuPro 110  
Db 307 TGGAGACAAGTGGGCTTTGGCCC-CCGACCCCTCTGCGTTCGTG-----TCC 353

QY 111 LeuAlaSerProGlyAlaProProAlaAlaProThrArgAlaSerProLeuGlyAlaArg 130  
Db 354 GAGCGCGCGGGGGCTCCCGAACTCCCTCGAAATCGTGGGCTCCATGTGGCTCCGG 413  
QY 131 -AlaSerProPro-----ArgSerGlyValSerLeuAlaArgProAlaPr 145  
Db 414 CAGCGTCCACCTCCCGCCACCTCGGGAAGGAGGGGTGGGAGTGCCTCCGCGCGTCC 473  
QY 145 oGlyCysProArgProAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysPr 165  
Db 474 CGGCTT----- 480  
QY 165 oGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 185  
Db 480 ----- 480  
QY 185 nGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly-----SerGlyLe 203  
Db 481 -----CCTCTTCCCGCCGACACCTCTCCGCGCGCGGTGGTGGCCGATCCGATT 533  
QY 203 uLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerValSerSerSerAl 223  
Db 534 GCTGTTCGAGCGCGAGTGGAGAGCGCCTGTGGAACAT----- 573  
QY 223 aThrAlaProSerSerValAlaAlaAlaThrSerGlyGlyArgProGlyLeuGlyAr 243  
Db 574 -----CG 575  
QY 243 gGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGlyIleTyrAl 263  
Db 576 AGGTGGAACAGTAAACAAAGGACTGCTCCGCGCGGTGGTGGCCGATCTATGC 635  
QY 263 aAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGluValGlnVa 283  
Db 636 AAATATGAGGATGTTTATATATATATATATATATATATATATATATATATATAT 695  
QY 283 lLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCysAspLeuVa 303  
Db 696 GAAAAATGAGGTATATATGAAGGAGTGTAAAACTTTACAGTCCGAGGTGATTGGT 755  
QY 303 lLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArgGluGlu 323  
Db 756 ACTTGATCCGCACATGAGAAAGTACAGATCCAGTTCGGGCGCGAAACGTGAAGAA 815  
QY 323 eMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLysAspMetAs 343  
Db 816 AATGAGAGTATTTTGTTCAAATGTTTACAGTTCAGCTTTTGTGTGATGATGGA 875  
QY 343 pSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLysValAsnGl 363  
Db 876 CTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGA 935  
QY 363 yGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGlyLeuThrAlaAsnGlu 383  
Db 936 CGAACACAAAGAGAGGACCTGAGCCCTGGGATGCGAGTGAACCTCACAGCCAATC 995  
QY 383 uLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMetPheArgTy 403  
Db 996 ACTTGAGGCTTTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCC 1055  
QY 403 rAsnGluGluAenTyrGlyValValSerThrTyrAspSerSerLeuSerSerTyrThrVa 423  
Db 1056 TAATGAAGAAAAATTTAGGTGTAGTGTACGTATGATGATGATGATGATGATGATG 1115  
QY 423 lProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAlaAsnGlnLe 443  
Db 1116 GCCCTTAGAAAGAGATTAATCAGAGAAATTTTAAACCGGAGGAGGCGGCAACAGTT 1175  
QY 443 uAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspAs 463  
Db 1176 AGCAGAGAAGAAATGAGTCAAGTCCAGGTACAAAGCTCAGTGGCGCTCGAAATCATGA 1235  
QY 463 pArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArgGlyHi 483

```

Db      1236 TAGGAGTGGGAAGAAATAACACAGAGCTTCAGAGAAATCCAGTGAACGTGAGGGGCA 1295
Qy      sSerIleAenThrArgGluAsnLysTyfileProProGlyGlnArgAsnArgGluValI1 503
Db      1296 CAGCATAAACACTAGGAGAAATAATATATCTCTCTGGACAAAGAAATAGAGAAGTCAT 1355
Qy      eSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySerGlySerMe 523
Db      1356 ATCTCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGCCAGCCCTGGATCGGGTCCAT 1415
Qy      pProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnAr 543
Db      1416 GCATCAAGATCCACTTCTCACACTTCAGATTCAACCCGAATTCGTGTTCAAGACCAAG 1475
Qy      gValValAenGlyGlyValProTrpProSerProCysProSerProSerProSerArgProPr 563
Db      1476 AGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCCATCTCCTTCTCTCGCCCAACC 1535
Qy      oSerArgTyGlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrArgPr 583
Db      1536 TTCTCGCTACCACTAGTCAGGTCCCAACTCTCTCCACTCGGCAGCCACCCCTACACGGCC 1595
Qy      oProSerArgProProSerArgProSerArgProProSerHisProSerAlaHisGlySe 603
Db      1596 GCCCTTCGGGGCCCCCTCGCGGCCATCCAGACCCCGCTCTCACCCCTCTGCTCATGGTTC 1655
Qy      rProAlaProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSe 623
Db      1656 TCCAGCTCTGCTCTACTATGCTTAACGATGCTTTCAGNAGGCGCTCCNAGGATGTC 1715
Qy      rProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSe 643
Db      1716 CCCAAAGGCCAGCGACATCTCGAAATCACAGAGTTTCTGCTGGGAGGGTTCCATATC 1775
Qy      rSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrProProValAl 663
Db      1776 CAGTGGCCCTAGAAATTTGATATCCACACCCACCCAGTGAACGAGCTACTCTCTCCAGTAGC 1835
Qy      aArgThrSerProSerGlyGlyThrTrpSerSerValValSerGly 678
Db      1836 AAGGACCACTCCCTCGGGGGNACGTGTGTCATCAGTGTGAGTGGG 1881

RESULT 14
AR447281
LOCUS      AR447281                1257 bp      mRNA      linear      PAT 20-FEB-2004
DEFINITION      Sequence 4 from patent US 6673535.
ACCESSION      AR447281
VERSION      AR447281.1      GI:42675576
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 1257)
AUTHORS      Pulst,S.M.
TITLE      Methods of detecting spinocerebellar Ataxia-2 nucleic acids
JOURNAL      Patent: US 6673535-A 4 06-JAN-2004;
            Cedars-Sinai Medical Center; Los Angeles, CA
FEATURES
            source
            1..1257
            /organism="unknown"
            /mol_type="mRNA"

ORIGIN
Alignment Scores:
Pred. No.:      3 67e-21      Length:      1257
Score:          1876.00      Matches:      365
Percent Similarity: 89.77%      Conservative: 21
Best Local Similarity: 84.88%      Mismatches: 18
Query Match:      26.95%      Indels:      26
DB:              6              Gaps:      5

us-10-802-228-2 (1-1312) x AR447281 (1-1257)

```

```

Qy      GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 176
Db      8 GGGCCGCTCACCATGTGCTGAAGCCGAGCGCAG----- 43
Qy      GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db      44 -----CCGCCCGCGCGCCGCACTGGCGGC 67
Qy      LysProGlyGlySerGlyLeuAlaSerProAlaAlaAlaProSerProSerSerSer 216
Db      68 AAGCCCGCGCGC---GGCCTGCTCTCGTCGCCCGCGCGCGCGCGC---TCGCCCGCG 12;
Qy      217 SerValSerSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSer--- 234
Db      122 GTGACCTCGGCTTCGCTGGTCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 181
Qy      235 ---GlyGlyGlyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
Db      182 GCGGCGCGCGCGCGCTCCCGGCTGGCGAGAGTCCGAACAGTAGCAAGAGGACTGCCTCAG 241
Qy      254 SerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSer 273
Db      242 CCTACGATTTCTTTTGATGGAATCTATGCAAACTGAGGATGGTTTCATATCTACGTCA 301
Qy      274 ValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPhe 293
Db      302 GTTGTGGATCGAAATGTGAAGTACAAAGTGAACGAGGAGCATATATGAAGAGATTTTT 361
Qy      294 LysThrTySerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGlu 313
Db      362 AAAACATACAGTCTCTAAGTGTGACTTGTGTTGATGTCGCACATGAGAAAAGTACAGAA 421
Qy      314 SerSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333
Db      422 TCCAGTTCGGGGCCCAAAACGTGAAGAAATAATGAGAGATGTTTGTTCAAATGCTCAGAC 481
Qy      334 PheValValValGlnPheLysAspMetAspSerSerTyralAlaLysArgAspAlaPheThr 353
Db      482 TTGCTGTGTGTACAGTTTAAAGATACAGACTCCAGTTTATGCACGAGAGATGCTTTTACT 541
Qy      354 AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrp 373
Db      542 GACTCTGCTCTCAGCGCAAGGTGAATGCTGAGCACAAGGAGAGGACCTCGAGCCCTTGG 601
Qy      374 AspAlaGlyGluLeuThrAlaAsnGluLuluLeuGluAlaLeuGluAsnAspValSerAsn 393
Db      602 GATCGAGGGAGCTCACCGCCAGCGAGAGCTGGAG---CTGGAGAAATGATGTGCTTAAT 658
Qy      394 GlyTrpAspProAsnAspMetPheArgTyraAsnGluGluAsnTyrglyValIleSerThr 413
Db      659 GGATGGGACCCCAATGACATGTTTCGATATAATATCAAGAGAATTTATGTTGCTGCTCCACA 718
Qy      414 TyrAspSerSerLeuSerSerTyThrValProLeuGluArgAspAsnSerGluGluPhe 433
Db      719 TATGATAGCAGTTTATCTTATATACGGTTCCTTTAGAAAGGACAACTCAGAAAGATTT 778
Qy      434 LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTy 453
Db      779 CTTAAACGGGAGGCAAGGGCAACAGTTAGCAAGAAATTTGAATCAAGTCTCAGTAC 838
Qy      454 LysAlaArgValAlaLeuGluAsnAspArgSerGluGluGluLysTyThrAlaVal 473
Db      839 AAAGCTCGTGTGCGCCCTTGAGAAATGATGACCGGAGTGAAGGAGGAGGAGGAGGAGG 898
Qy      474 GlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTy 493
Db      899 CAGAGAAATCTCAGTGAACCGGAGGCGCATGGCCCAACACTAGGAGCAATAAATATATT 958
Qy      494 ProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerPro 513
Db      959 CCTCTCGACAAAGAAACAGAGAAAGTCTCTATCTCTGGGAAAGTGGGAGAGAGAGTCA 1018

```



QY 514 ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp 533  
 Db 1019 CGGATGGGCGCCAGCTGGCCAGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGAT 1078  
 QY 534 PheAsnProAsnSerGlySerGlnArgValValAsnGlyGlyValProTrpProSer 553  
 Db 1079 TTCACCCGACGCTGGCTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCG 1138  
 QY 554 ProCysProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeu 573  
 Db 1139 CCTTGCCCATCTCTCTCGCCACCTTCTCGCTACAGTCAGGTCCCAACTCTCTT 1198  
 QY 574 ProProArgAlaAlaThrProThrArgPro 583  
 Db 1199 CCACCTCGGCGAGCCACCTTACACGGCT 1228

RESULT 15  
 AR632110  
 LOCUS AR632110 1257 bp mRNA linear PAT 14-FEB-2005  
 DEFINITION Sequence 19 from patent US 6844431.  
 ACCESSION AR632110  
 VERSION AR632110.1 GI:59773847  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCES  
 1 (bases 1 to 1257)  
 Pulet,S.M.  
 Nucleic acid encoding spinocerebellar ataxia-2 and products related thereto  
 JOURNAL Patent: US 6844431-A 19 JAN-2005;  
 Cedars-Sinai Medical Center; Los Angeles, CA  
 FEATURES  
 source 1..1257  
 /organism="unknown"  
 /mol\_type="mRNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,67e-21 Length: 1257  
 Score: 1876.00 Matches: 365  
 Percent Similarity: 89.77% Conservative: 21  
 Best Local Similarity: 84.88% Mismatches: 18  
 Query Match: 26.95% Indels: 26  
 DB: 6 Gaps: 5

US-10-802-228-2 (1-1312) x AR632110 (1-1257)

QY 157 GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 176  
 Db 8 GGGCGGCTCACCATGTGCTGTAAGCGCGAGCCGAG----- 43  
 QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196  
 Db 44 -----CCGCCCGCGCCCGCCATCGCGCG 67  
 QY 197 LysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaProSerSerSer 216  
 Db 68 AAGCCCGCGCGC---GGCTTGTCTCTGTCGCCGCGCGCGCGCGCGC---TCGGCGCG 121  
 QY 217 SerValSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSer----- 234  
 Db 122 GTGACCTCGGCTTCGTTGGTGGCGCGCGCGCGCGCGCGCGCTTCTCTCGCGCGCG 181  
 QY 235 ---GlyGlyGlyArgProGlyLeuGlyArgGlyArgGlySerAsnLysGlyLeuProGln 253  
 Db 182 GCGGCGCGCGCGCGCTCCCGCGCTGGCAGAGGTGGAAACAGTAGTCAAAAGGACTGCCTCAG 241  
 QY 254 SerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSer 273  
 Db 242 CCTACGATTTCTTTTGTAGTGAATCTATGCAACGTGAGGATGGTTCTATATACTTACGTCA 301  
 QY 274 ValValGlySerTyrCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPhe 293

Db 302 GTTGTGGATCGAAATGTGAAGTACAAGTGAAGAAACCGAGGCATATATGAAGGAGTTT 361  
 QY 294 LysThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGlu 313  
 Db 362 AAAACATACAGTCCTTAAGTGTGACTTGTGACTTGTGACTGACATGAGAAAAGTACAGAA 421  
 QY 314 SerSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333  
 Db 422 TCCAGTTCCGGGCCAAAACGTGAAGAAATATGAGAGTGTGTTGTTCAAATGCTCAGAC 481  
 QY 334 PheValValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr 353  
 Db 482 TTCGTTGTGTACAGTTTAAAGATACAGACTCCAGCTTATGACCGAGAGATGCTTTTACT 541  
 QY 354 AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrp 373  
 Db 542 GACTCTGCTCTCAGCGCAAGGTGAATGTTGAGCAACAAGGAGAGGACCTCGAGGCCCTGG 601  
 QY 374 AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsn 393  
 Db 602 GATGAGGGGAGCTCACGCCGAGCGAGGAGCTGGAG---CTGGAGAATGATGTCTTAAT 558  
 QY 394 GlyTrpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr 413  
 Db 659 GGAATGGACCCCAATGACATGTTTCGATATATGAAGAGAATATGTTGTTGTTGCCACA 718  
 QY 414 TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe 433  
 Db 719 TATGATAGCAGTTTATCTTCATATACGGTTCCTTTAGAAAAGGACAACTCAGAAGAAAT 778  
 QY 434 LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluLysSerAlaGlnTyr 453  
 Db 779 CTTAAACGGGAGCGCAAGGCAACCGATTAGCAGAGAAATTTGAATCCAGTCTCAGTAC 838  
 QY 454 LysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluGlyLysTyrThrAlaVal 473  
 Db 839 AAGGCTGCTGTCGCCCTTGAGAAATGATGACCGGAGTGAGAAAGAAATACACAGCAGTC 898  
 QY 474 GlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIle 493  
 Db 899 CAGAGAAACTGCAGTGCACCGGAGGGGCATGCCCCAACACTAGGACAAATAATATATT 958  
 QY 494 ProProGlyGlnArgAsnArgGluValIleSerTrpGlySerClyArgGlnAsnSerPro 513  
 Db 959 CCTCTCTGGACAAAGAAACAGAGAAGTCTTATCCTGGGGAAGTGGGAGACAGACTCACCA 1018  
 QY 514 ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp 533  
 Db 1019 CGGATGGGCGAGCTGGCCAGGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGAT 1078  
 QY 534 PheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTrpProSer 553  
 Db 1079 TTCACCCGACGCTGGCTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCG 1138  
 QY 554 ProCysProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeu 573  
 Db 1139 CCTTGCCCATCTCTCTCTCGCCACCTTCTCGCTACAGTCAGGTCCCAACTCTCTT 1198  
 QY 574 ProProArgAlaAlaThrProThrArgPro 583  
 Db 1199 CCACCTCGGCGAGCCACCTTACACGGCT 1228

Search completed: December 23, 2005, 10:45:44  
 Job time : 14148 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 23, 2005, 00:44:12 ; Search time 1532 Seconds

(without alignments)  
5707.622 Million cell updates/sec

Title: US-10-802-228-2

Perfect score: 6961

Sequence: 1 MRSAAAPSPAVATESRRF.....TAHPYMTSPVQAHHQOOL 1312

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=slp  
-Q=/cgn2/USPTO.spool\_p/US10802228/runat\_21122005\_110224\_9692/app\_query.fasta\_1.1479  
-DB=N Geneseq -QMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10802228@cgn 1.1.1232 @runat\_21122005\_110224\_9692 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 21.\*

1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: geneseqn2004as.\*  
13: geneseqn2004bs.\*  
14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	2	AAV06552 Human SCA
2	6961	100.0	4481	2	Aaz23428 Human SCA
3	6961	100.0	4481	10	ADD18753 Human dis
4	6961	100.0	4481	10	ADH69466 Human spi

5	6961	100.0	4481	10	ABV75291 Human SCA
6	6961	100.0	4481	14	ADZ49164 Insulin S
7	6950.5	99.8	4367	2	AAV30270 Gene caus
8	6377	91.6	4200	2	AAT78912 Spinocere
9	6025.5	86.6	4225	10	ABV75290 Mouse SCA
10	5606.5	80.5	3798	2	AAV06553 Mouse SCA
11	2463	35.4	2006	3	AAf21628 Human bre
12	1876	27.0	1257	10	ABV75289 Mouse SCA
13	1816.5	26.1	4674	2	AAx90421 Human ata
14	1816.5	26.1	4674	8	ABZ72015 Human chr
15	1816.5	26.1	4674	10	ADH23078 Human chr
16	1257	18.1	4369	13	ADR08336 Full leng
17	1042.5	15.0	1707	6	ABK34831 Human cdn
18	1012.5	14.5	623	2	AAV7229 SCA2 gene
19	857	12.3	516	2	AAV06551 SCA2 gene
20	826	11.9	1226	6	ABK84413 Human cdn
21	799	11.5	485	9	ACH41056 Human foe
22	787	11.3	459	10	ADH69479 Human SCA
23	735	10.6	3210	4	ABL08977 Drosophil
24	725	10.4	5375	4	ABL08976 Drosophil
25	643	9.2	355	2	AAV17224 SCA2 gene
26	598	8.6	93631	11	ACN43984_4 Continuation (5 of
27	587	8.4	13416	10	AAZ55815 Micromono
28	587	8.4	60196	10	AAZ55810 Micromono
29	568.5	8.2	6492	12	ADK16023 Streptomy
30	568	8.2	30690	3	AAA92301 S. avermi
31	568	8.2	30690	4	AAH79277 Streptomy
32	566	8.1	11238	10	AAZ55817 Micromono
33	564	8.1	88400	13	ADU47667 S. cyaneo
34	556.5	8.0	138203	14	ADZ51725 FR-008 po
35	543.5	7.8	85915	13	ADL15447 Streptomy
36	541.5	7.8	67167	14	ADY80053 Nucleotid
37	541	7.8	29467	14	ADY80051 Contig #1
38	535.5	7.7	13987	2	AAH80415 Hybrid sr
39	535.5	7.7	44377	2	AAT80414 Platenoli
40	535.5	7.7	44377	2	AAT78508 Platenoli
41	529	7.6	104096	13	ADX56092 Streptomy
42	528	7.6	65140	4	AAAD17184 Streptomy
43	528	7.6	125401	4	AAAD17186 Streptomy
44	520.5	7.5	73995	11	ACN43986 Human gen
45	520.5	7.5	84428	12	ADM45913 Streptomy

ALIGNMENTS

RESULT 1  
AAV06552  
ID AAV06552 standard; cdna; 4481 BP.  
XX  
AC AAV06552;  
XX  
DT 06-JUL-1998 (first entry)  
XX  
DE Human SCA2 cdna including CAG repeat region.  
XX  
KW SCA2 gene; spinocerebellar ataxia-2; ataxin-2; human; diagnosis;  
KW olivoponto-cerebellar atrophy; ss; ds.  
XX Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT CDS 164..4101  
FT /\*tag= a  
FT primer\_bind complement(631..648)  
FT /\*tag= b  
FT repeat\_region /note= "primer SCA2-A binding site"  
FT 658..723  
FT /\*tag= f  
FT repeat\_unit /note= "CAG repeat region"  
FT 658..660  
FT /\*tag= g  
FT repeat\_unit /note= "CAG repeat"  
FT 661..663

```
FT FT /*tag= h
FT FT /note= "CAG repeat"
FT FT 664..666
FT FT /*tag= i
FT FT /note= "CAG repeat"
FT FT 667..669
FT FT /*tag= j
FT FT /note= "CAG repeat"
FT FT 670..672
FT FT /*tag= k
FT FT /note= "CAG repeat"
FT FT 673..675
FT FT /*tag= l
FT FT /note= "CAG repeat"
FT FT 676..678
FT FT /*tag= m
FT FT /note= "CAG repeat"
FT FT 679..681
FT FT /*tag= n
FT FT /note= "CAG repeat"
FT FT 685..687
FT FT /*tag= o
FT FT /note= "CAG repeat"
FT FT 688..690
FT FT /*tag= p
FT FT /note= "CAG repeat"
FT FT 691..693
FT FT /*tag= q
FT FT /note= "CAG repeat"
FT FT 694..696
FT FT /*tag= r
FT FT /note= "CAG repeat"
FT FT 700..702
FT FT /*tag= s
FT FT /note= "CAG repeat"
FT FT 703..705
FT FT /*tag= t
FT FT /note= "CAG repeat"
FT FT 706..708
FT FT /*tag= u
FT FT /note= "CAG repeat"
FT FT 709..711
FT FT /*tag= v
FT FT /note= "CAG repeat"
FT FT 712..714
FT FT /*tag= w
FT FT /note= "CAG repeat"
FT FT 715..717
FT FT /*tag= x
FT FT /note= "CAG repeat"
FT FT 718..720
FT FT /*tag= y
FT FT /note= "CAG repeat"
FT FT 721..723
FT FT /*tag= z
FT FT /note= "CAG repeat"
FT FT 740..757
FT FT /*tag= c
FT FT /note= "primer SCA2-B binding site"
FT FT 899..900
FT FT /*tag= e
FT FT /note= "predicted splice site"
FT FT 1070..1091
FT FT /*tag= d
FT FT /note= "primer SCA2-14B binding site"
```

```
PR 19-JUL-1996; 96US-0022207P.
PR 08-OCT-1996; 96US-00727084.
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX Pulst S;
XX
XX WPI; 1998-086523/08.
DR P-PSDB; AAW33807.
XX
XX Nucleic acids encoding human and mouse ataxin 2 - a product of the
FT spinocerebellar ataxia 2 gene, SCA2; useful in the diagnosis of ataxia
FT type 2.
XX
XX Claim 6; Page 52-58; 98pp; English.
XX
XX This cDNA sequence corresponds to a novel SCA2 gene encoding a human
CC spinocerebellar ataxin-2 (SCA2) polypeptide, designated ataxin-2 (see
CC AAW33807). A trisomy 21 foetal brain cDNA library and an adult human
CC frontal cortex cDNA library in lambda ZapII were screened with probes
CC obtained by PCR amplification of plasmid AAP65122B (see AAV06551). PCR
CC products were used to screen the human adult frontal cortex library, and
CC 5' clones were obtained by RT-PCR of placental mRNAs. Overlapping clones
CC was used to generate the composite 481 bp sequence. Ataxin type 2 can be
CC diagnosed by detecting a genomic or transcribed mRNA sequence in an
CC individual having an expanded CAG repeat at a location corresponding to
CC the CAG repeat region of the SCA2 gene. The presence of at least 13 CAG
CC repeats above the normal level (22, occasionally 23, repeats) is
CC indicative of SCA2. Primers (see AAT99640-41) amplifying at least this
CC region are used for diagnosis. Also claimed are kits for detecting
CC mutations at the SCA2 locus, antisense oligonucleotides, and transgenic
CC animals useful for studying the physiological roles of ataxin-2 and its
CC effect upon behaviour
XX
XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 5.27e-200 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-802-228-2 (1-1312) x AAV06552 (1-4481)
Qy 1 MetArgSerAlaAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGCGCTCAGCGGCGGCAGCTCCTCGGAGTCCCGGGTGGCCACGAGTCTCGCGCGCTTC 222
Qy 21 AlaAlaAlaArgTtpProGlyTtpArgSerLeuGlnArgProAlaAlaArgSerGlyArg 40
Db 223 GCCGAGCCAGGTGGCCCGGGTGGCGTCCGTCAGCGCGCGCGCGGAGGAGGGCGG 282
Qy 41 GlyGlyGlyAlaAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 283 GCGGCGGTGGCGCGCGCGCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGCC 342
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 343 GGCCCCCTCCCTCCCGCAGAGCTCGCTCCCTCCGCGCTCAGACTGTGTTGGTAGCAAC 402
Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgLeuLeuLeuGlyGlyGlyPro 100
Db 403 GGCACCGCGCGCGCGCGGTTCGGCCCGGCTCCCGCGGCGCTCTTGGTCTCGGGGGGCT 462
Qy 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
Db 463 CCGCGCCCTTCGTCGTCGTCCTTCTCCCTCCCTCCGCGAGCCCGGGCGCCCTCCGCGCGG 522
Qy 121 ProThrArgAlaSerProLeuGlyAlaAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 523 CCAACCGCGCGCTCCCGCTCGGCGCGCGTCCGTCGCCCGCGCGCTTCGCGGCTCTCCTTG 582
```

QY	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160
DB	583	GGCGCCCGGCTCCCGGCTGTCCCGCGCGGTGCGAGCGGTGTATGGGCGCCCTCACC	642
QY	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180
DB	643	ATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702
QY	181	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
DB	703	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	762
QY	201	SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSerSer	220
DB	763	AGCGGCTTCTAGCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	822
QY	221	SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly	240
DB	823	TCCTCGGCCACGGCTCCCTCCTCGGTGTCGCGCGACCTCCGCGCGCGCGCGCGCG	882
QY	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260
DB	883	CTGGCGCAGAGGTGCAACAACTAACAAAGACTGCGCTCAGTCTACGATTTCTTTT	942
QY	261	IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280
DB	943	ATCTATGCAAAATATGAGGATGTTTCATATACATCATGTTGTTGGCTCCAAATGTGA	1002
QY	281	ValGlnValLysAsnGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300
DB	1003	GTACAGTGAAAAATGAGGATATATATGAAGGAGTTTTTAAAACTTACAGTCCGAAGT	1062
QY	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerGlyProLysArg	320
DB	1063	GATTTGGTACTTGATCCCGCACATGAGAAAGTACAGATCCAGTTCCGGGCGCAACGT	1122
QY	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys	340
DB	1123	GAAGAAATAATGAGAGATTTTTTGTTCAAATGTTTCAGACTTTTGTGTACAGTTTAAA	1182
QY	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
DB	1183	GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCAA	1242
QY	361	ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla	380
DB	1243	GTGAATGGCGAACAACAGAGAGGACCTGGAGCCCTGGATGCAGGTGAACCTCACAGCC	1302
QY	381	AsnGluGluLeuAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet	400
DB	1303	AATGAGGAACTTGAGGCTTTTGGAAATGACGTATCTAAATGGATGGGATCCCAATGATG	1362
QY	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420
DB	1363	TTTCGATATAATGAAGAAATTTATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTCG	1422
QY	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440
DB	1423	TATACAGTGCCTTTAGAAAGATTAACCTCAAGAAATTTTTTAAACCGGAGAGCAGGGCA	1482
QY	441	AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460
DB	1483	AACCAGTTACAGAGAAATTTAGTCAAGTGCCAGTACAAAGCTCGAGTGGCCCTGGAA	1542
QY	461	AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg	480
DB	1543	AATGATGATAGGATGAGGAGAGAAATAACACAGCAGTTCCAGAGAAATTCAGTGAACGT	1602
QY	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500
DB	1603	GAGGGGCACAGCATAAACACTAGGAAAAATAAATATATTTCTCTCTGACAAAGAAATAGA	1662

QY	501	GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520
DB	1663	GAAGTCATATCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGGCGAGCTGGATCG	1722
QY	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540
DB	1723	GGCTCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCACCCGAATTCGTGTTCA	1782
QY	541	AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer	560
DB	1783	GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCTCTCTCT	1842
QY	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro	580
DB	1843	CGCCACACCTTCTCGGTACAGTCAGTCCCACTCTCTTCCACCTCGGCGACGCCCTT	1902
QY	581	ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla	600
DB	1903	ACACGGCGCCCTCCAGCCCGCTCGGGCCATCCAGACCCCGCTCTCACCCCTCTGCT	1962
QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
DB	1963	CATGGTCTCCAGCTCCTGTCTACTATGCTTAAACGATCTCTTCAAGAGGCGCTCCA	2022
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
DB	2023	AGGATGTCCCAAGAGCCCGCAGCATCTCTGAAATCACAGAGTTTCTCTGGAGGGGT	2082
QY	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro	660
DB	2083	TCATATCCAGTGGCTAGAAATTTGTATCCACACCCACCCAGTGAAGCAGCTACTCCT	2142
QY	661	ProValAlaArgThrSerProSerGlyThrTyrSerSerValValSerGlyValPro	680
DB	2143	CCAGTAGCAAGACACAGTCCCTCGGGGAAACGTGTCTCATCTAGTGTGCTAGTGGGTCCA	2202
QY	681	ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr	700
DB	2203	AGATTATCCCTTAAACACTCATGACCCAGGTCTCCAGACAGAACAGTATTGGAATACC	2262
QY	701	ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla	720
DB	2263	CCAGTGGCGCAGTCTTGTCTCTCCCAAGCTGGTATTATTCCAAGTGAAGCTGTGCC	2322
QY	721	MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr	740
DB	2323	ATGCCTATTCAGCTGCATCTCCTACGCTGCTAGTCTCTGCATCGAACAGAGCTGTACC	2382
QY	741	ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly	760
DB	2383	CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCGAGAACTCTCTGCAGGG	2442
QY	761	AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys	780
DB	2443	AATAAGAAAAATTTAAACCAATGAACATCACCTAGTCTCTCAAAAGCTGAAACAAA	2502
QY	781	GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys	800
DB	2503	GGTATATCACAGTGTCTTCTGAACATAGAAAAACAGATTGATTTAAAGAAATTTAAG	2562
QY	801	AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys	820
DB	2563	AATGATTTAGGTTACAGCCAACTTCTACTTCTGAATCTATGATCAACTTACTTAAACAAA	2622
QY	821	AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys	840
DB	2623	AATAGAGAGGAGAAAAATCAAGAGATTGTATCAAGACAAAAATTTGACCAAGTGTCTAAG	2682
QY	841	AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer	860
DB	2683	GATTCCTTTCATTGAAATAGCAGCAGCACTGTACCAGTGGCAGCAGCAAGCCGAATAGC	2742
QY	861	ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr	880



CC the severity of a pathological condition mediated by upregulated cell  
 CC proliferation or cell survival e.g. neoplastic, malignant, autoimmune or  
 CC fibrotic conditions. This sequence encodes the human SCA2 polypeptide  
 CC described in the method of the invention

XX  
 SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 5,27e-200 Length: 4481  
 Score: 6961.00 Matches: 1312  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-802-228-2 (1-1312) x AA223428 (1-4481)

QY	1	MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgPhe	20
DB	163	ATGCGCTCAGCGCGCGAGCTCTCGAGTCCGCGGTGCGCCACCGAGTCTCGCCGCTTC	222
QY	21	AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg	40
DB	223	GCAGCGAGCAGGTGGCCCGGGTGGCGCTCGCTCCAGCGCGCGCGCGGCGGCGG	282
QY	41	GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro	60
DB	283	GGCGCGGTGGCGCGCGCGCGAGCTATCCCTCCGCGCGCGCTCCCGCGCGCGCC	342
QY	61	GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn	80
DB	343	GGCGCGCGCTCCCTCCGCGCAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGTGTAGCAAC	402
QY	81	GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro	100
DB	403	GGCAACGGCGCGCGCGGTTCGGCCCGCGCTCCCGCGCGCTCTTGGTCTCGCGCGCT	462
QY	101	ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProAlaAla	120
DB	463	CCCGCGCGCTTCGTCTCGTCTCTTCTCCCGCTCCCGCGCGCGCGCTCCCGCGCGCG	522
QY	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu	140
DB	523	CCAAACCGCGCGCTCCCGCTCGCGCGCGCGTTCGCTCCCGCGCGCTTCGCGCGCT	582
QY	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160
DB	583	GCAGCGCGCGCTCCCGCTGTCCCGCGCGCGGTGCGAGCGGTGTATGGCGCGCTCAC	642
QY	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180
DB	643	ATGTCGCTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702
QY	181	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
DB	703	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	762
QY	201	SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerValSerSer	220
DB	763	AGCGCGCTTCCTAGCGTCCCGCGCGCGCGCTTCGCGCGCTTCGCTCGCTCTCGCT	822
QY	221	SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly	240
DB	823	TCCTCGCGCGCGCTCCCTCTCGGTGTCTCGCGCGCGCTCCCGCGCGCGCGCGCG	882
QY	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260
DB	883	CTGGCGCAGGTCCGAAACAGTAACAAAGAGCTGCTCTCAGTCTACGATTTCTTTGTGGA	942
QY	261	IleTyrAlaAsnMetArgMetValIleIleLeuThrSerValValGlySerLysCysGlu	280
DB	943	ATCTATGCAATATGAGATGGTTTCATATATCTTACATCAGTTGTGTGGCTCCAAATGTGAA	1002

QY	281	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300
DB	1003	GTACAAGTGCAAAATGGAGGTATATATGAAGAGATTTTAAACCTTTACAGTCCGAAGTGT	1062
QY	301	AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg	320
DB	1063	GATTTGGTACTTGTGCGCACATGAGAAAGTCAGAAATCCAGTTCGGGGCGGAAACGT	1122
QY	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys	340
DB	1123	GAAGAAATAATGGAGAGTATTTTGTTCAAATGTTTCAGACTTTGTTGTGTACAGTTTAA	1182
QY	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
DB	1183	GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTCTGACTCTGCTATCAGTGTCTAAA	1242
QY	361	ValAsnGlyGluHisLysGlyLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla	380
DB	1243	GTGAATGGCGAACACAAAGAGAGACCTGGAGCCCTGGGATGCGAGTGAACCTCACAGCC	1302
QY	381	AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet	400
DB	1303	AATGAGAACTTTCAGGCTTTGGAAAATGACGTATCTAATGGATGGGATCCCAATGATATG	1362
QY	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420
DB	1363	TTTCGATATAATGAAGAAATTTATGTTGTGTGTCTTACGATGATAGCAGTTTATCTTCG	1422
QY	421	TyrThrValProLeuGluArgAspAsnSerGluPheLeuLysArgGluAlaArgAla	440
DB	1423	TATACAGTCCCTTAGAAGAGATAACTCAGAAGAAATTTTAAACCGGAGCAAGCGCA	1482
QY	441	AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460
DB	1483	AACCACTAGCAGAGAAATTTAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTCGAA	1542
QY	461	AsnAspAspArgSerGluGluGlyTyrThrAlaValGlnArgAsnSerSerGluArg	480
DB	1543	AATGATGATAGAGTGAAGAAATAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT	1602
QY	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500
DB	1603	GAGGGCGCAGCATAAACACTAGGGAATAATAATATATTTCTCTCGACAAAGAAATAGA	1662
QY	501	GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520
DB	1663	GAAATCATATCTCTGGGAAAGTGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCG	1722
QY	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540
DB	1723	GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCACCCGAAATTCGGTTCA	1782
QY	541	AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer	560
DB	1783	GACCAAGAGTAGTTAATGAGGTGTTCCTTGCCCATCGCTTGCCTATCTCTCTCTCT	1842
QY	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro	580
DB	1843	CGCCCCCTTCTCGCTACCACTCAGGTCCCACTCTCTTCCACTCGGCGAGCCACCCCT	1902
QY	581	ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla	600
DB	1903	ACACGGCGCGCTCCAGGCGCCCTCGCGGCCATCCAGACCCCGCTCTCACCCCTCTGCT	1962
QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
DB	1963	CATGTTTCTCAGCTCTGCTCTACTATGCTTAAACGATGCTTTCAGAGGCGCTCCA	2022
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
DB	2023	AGGATGTCCCAAGAGCGGCGACATCTCTCGAAATCACAGAGTTTCTGCTGGAGGGT	2082
QY	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro	660



```
Db 2083 TCCATATCAGTGGCGCTAGAAATTTGATATCCCAACACCCAGCCAGTGAAGCAGCTACTCCT 2142
Qy 661 ProValAlaArgThrSerProSerGlyThrTrpSerSerValValSerGlyValPro 680
Db 2143 CCAGTAGCAAGACACAGTCCCTCGGGGGAACGTGGTCATCAGTGGTCAAGTGGGTTC 2202
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTATCCCTTAAACCTCATAGACCCAGGTCTCCAGACAGAACAGTATTGGAAATACC 2262
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db 2263 CCAGTGGCCAGTCTTCTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGCC 2322
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Db 2323 ATGCCTATTCCAGCTGCATCTCTAGCCCTGCTAGTCTGTCATCGAACAGAGCTGTACC 2382
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Db 2383 CCTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGCAGAACTCTCTCGAGGG 2442
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db 2443 AATAAGAAAAATTAAACCCAAATGAACATCACCTAGCTTCTCAAAAGCTGAACACAA 2502
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
Db 2503 GGTATATCACCAAGTGTCTTCTGAAATAGAAACAGATTGATTTAAAGNAATTTAAG 2562
Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db 2563 AATGATTTTAGGTTCAGCCAAAGTTCTACTTCTGAATCTATGCATCAACTACTAAACAA 2622
Qy 821 AsnArgGluGlyGluLysSerArgAspIleLysAspLysIleGluProSerAlaLys 840
Db 2623 AATAGAGGGGAGAAAAATCAAGAGATTGTATCAAAAGACAAAAATTTGAACCAAGTCTAAG 2682
Qy 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db 2683 GATCTCTTCATTGMAAATAGCAGCAGCAACTGTACCAAGTGGCAGCAGCAAGCCGAATAGC 2742
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCCAGCATTTCCCTTCAATACTTAGTAACAGGAGCAACAGAGGGGACCTCAGGTCACT 2802
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
Db 2803 TCCCAGGGGTTTCAGACTTCCAGCCCGAGCATGTAAACAGAGAAAGACGATTAAGGAAGAG 2862
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAAGCCAGCTCAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGAGGTTTC 2922
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCACGTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAA 2982
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyThrGlnPro 960
Db 2983 GCACAACTTAGGCCCATCTATGTGGGTTCATCAACAGCCCAACTCCAGTTTATATCTCAGCT 3042
Qy 961 ValCysPheAlaProAsnMetMetTyProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTTGTTTGCACCAAAATATGATGATATCCAGTCCAGTCCAGGCCAGCGGTGCAACTTTA 3102
Qy 981 TyProIleProMetThrProMetProValAsnGlnAlaLysThrTyArgAlaValPro 1000
Db 3103 TACCCAAATACCTATGACGCCCATGTCAGGTGAATCAAGCCCAAGACATATAGACAGTACCA 3162
Qy 1001 AsnMetProGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
```

```
Db 3163 AATATGCCCCAACAGCGGCAGACCGATCATCAGAGTCCATGATGCACCCAGCGTCA 3222
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyThrSerThrGlnTyValAlaTy 1040
Db 3223 GCAGCGGGCCACAGGATTGACGCCACCCACCCAGCTTACTCCAGCAATATGTTGCCTAC 3282
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAG 3342
Qy 1061 HisProHisValTySerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCTCATGTCTATAGTCTGTAAATACAGGGTAATGCTAGAAATGATGGCACCAACA 3402
Qy 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyGlyAlaHisGluGlnThr 1100
Db 3403 CACGCCACCGCTGCTTTAGTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3462
Qy 1101 HisAlaMetTyAlaCysProLysLeuProTyAsnLysGluThrSerProSerPheTy 1120
Db 3463 CATGGATGTATGATGTCTCCAAATTTACCATCAACAAGGAGACAAGCCCTTCTTTCTAC 3522
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGCCATTTCCAGGGCTCCCTTGCTCAGCAGATATGCGCACCTTAAGCTTACCCTGCAC 3582
Qy 1141 ProHisThrProHisProGlnProSerSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 3583 CCACATATCCACACCCCTCAGCTTACCTACCCCTCAGCAGCAGCAAGCAACAT 3642
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Db 3643 GGTGGAAGTCACTCCTGCACCCAGTCTCTTTCAGCACCATCAGCACCAGCGCCAGGCT 3702
Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCCAGTTCACAGCAGCAGTACCCATTTTACCAGCGGGCTTGCGCCAACT 3762
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3763 CCACCTCCATGACACCTGCTCCCAACAGCAGTCCGCCACAGAAATAGTTTCCAGCAGCA 3822
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyThrAsnProPro 1240
Db 3823 CAACAGACTGTCTTTAGCATCCATCTTCTCACCTTACCGCGCGGTATACCAACCCACCC 3882
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGGCCCACTACCTCAGGCTCATGTACAGTACAGAAATGGTTCTTCTCATCCAACT 3942
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGCGCCAATGATGCTAATGACGACACAGCCACCCCGCGGTCCCCAGCGCCCTC 4002
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyMetThr 1300
Db 4003 GCTCAAAGTGCACTACAGCCCATTTCCAGTCTCGAACACAGCGCATTTTCCCTATATGACG 4062
Qy 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
Db 4063 CACCTTCAGTACAAAGCCCAACCAACAGCAGTTG 4098
```

## RESULT 3

```
ADD18753
ID ADD18753 standard; DNA; 4481 BP.
XX
AC ADD18753;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID184.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
```

KW	antiarteriosclerotic; vulnary; gene therapy;
KW	hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW	inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW	glucose transportation; catecholamine synthesis; iron transport;
KW	nicric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW	retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW	inflammatory condition; wound healing; gene; ds.
XX	
OS	Homo sapiens.
OS	
PN	WO2003018621-A2.
XX	
PD	06-MAR-2003.
XX	
PF	23-AUG-2002; 2002WO-GB003892.
XX	
PR	23-AUG-2001; 2001GB-00020558.
PR	05-OCT-2001; 2001GB-00024037.
XX	
PA	(OXFO-) OXFORD BIOMEDICA UK LTD.
XX	
PI	Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX	
DR	WPI; 2003-290046/28.
DR	P-FSDB; ADD18752.
XX	
PT	New substantially purified polypeptide, useful for diagnosing or treating
PT	a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT	injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT	wound healing.
XX	
PS	Claim 27; SEQ ID NO 184; 424pp; English.
XX	
CC	This invention relates to novel human genes and gene product which are
CC	implicated in certain disease states. Compounds which modulate the
CC	proteins of the invention may have cytostatic, antiinflammatory,
CC	ophthalmological, antiarteriosclerotic or vulnary activities. The
CC	sequences of the invention may be useful for gene therapy. The invention
CC	may be useful for diagnosing or treating a hypoxia-regulated condition,
CC	such as tumorigenesis, angiogenesis, apoptosis, inflammation,
CC	erythropoiesis, or the biological response to hypoxia conditions
CC	including processes such as glycolysis, gluconeogenesis, glucose
CC	transportation, catecholamine synthesis, iron transport or nitric oxide
CC	synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC	injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC	inflammatory conditions or wound healing. The present sequence is that of
CC	a disease related protein encoding DNA sequence of the invention.
XX	
SQ	Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	5-27e-200 Length: 4481
Score:	6961.00 Matches: 1312
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	10 Gaps: 0
US-10-802-228-2 (1-1312) x ADD18753 (1-4481)	
Qy	1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db	163 ATGCGCTCAGCGCGCAGCTCTCGAGTCCCGGTGGCCACCGAGTCTCGCGCTTC 2222
Qy	21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db	223 GCGCAGCCAGGTGGCCCGGTGGCGCTCGCTCCAGCGCGCGCGCGAGCGGCGG 2822
Qy	41 GlyGlyGlyAlaAlaProGlyProTyProSerAlaAlaProProProGlyPro 60
Db	283 GCGCGCGGTGGCGCGCGCCCGGACCGATTCCTCGCGCGCCCTCCCGCCCGGCCCC 3422
Qy	61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80

Db 1423 TATACAGTGCCTTTAGAAAGACATAACTCAGACAGAAATTTTAAACGGGAAGCAAGGCCA 1482  
Qy 441 AenGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
Db 1483 AACCAAGTTAGCAGAAAGAAATTGAGTCAAGTGCCTAGTACAAAGCTCGAGTGGCCCTGGAA 1542  
Qy 461 AenAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
Db 1543 AATGATGATAGGAGTGAGAGAGAAATAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602  
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
Db 1603 GAGGGGCACAGCATAAACACTAGGGAATAATAATATATTCCTCTCGGACAAAGAAATAGA 1662  
Qy 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
Db 1663 GAAGTCATATCTCGGGAAGTGGAGACAGAATTACCCGCGTATGGGCCAGCCTGGATCG 1722  
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAAACCCGAATTTCTGGTTCA 1782  
Qy 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560  
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCTGGCCATCGCCTTGCCTCATCTCCTTCCT 1842  
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
Db 1843 CGCCCACTTCTCGCTACCAAGTCAGGTCCCAACTCTCTCCACCTCGGGCAGCCACCCCT 1902  
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
Db 1903 ACACGGCGCCCTCCAGGCCCCCTCGGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Db 1963 CATGTTCTCCAGCTCCTGCTCTACTATGCTTAAACGCATGTCTTCAGAAAGGGCCCTCA 2022  
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Db 2023 AGGATGTCCCCAAAGCCCGCAGCATCTCGAATATCACAGATTTCTGCTGGAGGGGT 2082  
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
Db 2083 TCCATATCCAGTGGCTAGAAATTTGATATCCCAACCCACCCAGTGAAGCAGCTACTCCT 2142  
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680  
Db 2143 CCAGTAGCAAGACCAGTCCCTCGGGGGGAACGTGGTCAATCAGTGGTCAAGTGGGGTTCCA 2202  
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
Db 2203 AGATTATCCCTTAAACTATAGACCCAGGCTCCACAGACAGACAGTATTTGGAATATCC 2262  
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720  
Db 2263 CCCAGTGGGCCAGTCTTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGCC 2322  
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
Db 2323 ATGCTATTTCCAGCTGCGATCTCTACGCTGTAGTCTGTGATCGAATCGAAACAGAGCTGTATCC 2382  
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
Db 2383 CTTCTAGTAGGCTAAAGATTTCCAGGCTTCAGATCAGAGCGAGAACTCTCCTCGCAGGG 2442  
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
Db 2443 AATAAAGAAAAATATTAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502  
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
Db 2503 GGTATATCACCAGTTGTTTCTGAACATAGAAACAGATTGATTTTAAAGAAATTTTAAG 2562

Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
Db 2563 AATGATTTTAGGTTTACAGCCAAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAA 2622  
Qy 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
Db 2623 AATAGAGAGGGAGAAAAATCAAGAGATTTGTATCAAAACACAAAATTTGAACCAAGTGTCTAAG 2682  
Qy 841 AspSerPheIleGluAsnSerSerAsnSerSerAsnSerSerGlySerSerLysProAsnSer 860  
Db 2683 GATTCTTTTCATTGAAATATAGCAGCAGCAACTGTACCAAGTGCACGACGACCCCAATAGC 2742  
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
Db 2743 CCCAGCATTTCCCTTCAATACTTAGTAACACGAGCACAAGAGGGGACCTGAGTCACT 2802  
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnLysAspAspLysGluGlu 900  
Db 2803 TCCCAAGGGGTTTCAGACTTCCAGCCAGCATGTAAACAAGAGAAAGACGATAAGGAAGAG 2862  
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
Db 2863 AAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCCAAAGAGATT 2922  
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
Db 2923 AACCCAGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCCTCAA 2982  
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
Db 2983 GCACAACTTAGCCCATCTATGGTGGGTCAACAAGCACAACCTCCAGTTTATACTCAGCCT 3042  
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
Db 3043 GTTTGTTTTGCACCAATATGATGATCCAGTCCAGTGCCAGGGCGTCAACCTTTA 3102  
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
Db 3103 TACCCAATACCTATGACGCCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCA 3162  
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020  
Db 3163 AATATGCCCAACACGGCGAAGACAGCATCATCAGAGTGCATGATGCACCCAGCGTCA 3222  
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
Db 3223 GCAGGGGGCCACCGATTGACGCCACCCACAGCTTACTCCAGCAATATGTGGCTTAC 3282  
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
Db 3283 AGTCCTCAGCAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAG 3342  
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
Db 3343 CATCCTCATGTCTATAGTCTCTGTAATACAGGGTAAATGCTAGAAATGATGGCACCAACA 3402  
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
Db 3403 CACGCCACCGCTGGTTTAGTATCTTTCACAACTCAGTACGGGGCTCATGACAGACG 3462  
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
Db 3463 CATGCGATGATGATGTGTCCCAATTACCATAACAAGGAGACAAGCCCTTCTTCTTAC 3522  
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
Db 3523 TTTGCCATTTCCACGGGCTCCCTTGCTCAGCAGTATGTCGCAACCCCTAAAGCTACCTGAC 3582  
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160  
Db 3583 CCACATACTCCACACCTTCAGCCTTCAGCTTACCCCTTGGACAGCAGCAAGCCACAT 3642

QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAla 1180  
DB 3643 GGTGAAGTCATCTGACCACTGCTGTTTCAGCACCATCAGCAGCGCCCGAGGCT 3702  
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
DB 3703 CTCATCTGGCAGTCCACAGCAGCAGTCCAGCATTTACACGCGGGCTTGGCCNACT 3762  
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
DB 3763 CCACCTCCATGACACCTGCTCCACACGCGAGTCGCCACAGATAGTTTCCAGCAGCA 3822  
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
DB 3823 CAACGACTGTCTTTACGATCCATCTCTCAGTTTCAGCGCGGTATACCAACCCACCC 3882  
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
DB 3883 CACATGGCCCGTACCTCAGGCTCATGTACAGTCAGGATGTTCTTCTCATCCNACT 3942  
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
DB 3943 GCCCATGCGCCATGATGCTAATGACGACACAGCCACCGCGGTCCCGAGCGCGCTC 4002  
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
DB 4003 GCTCAAAGTGCACTACAGGCCATTTCCAGTCTCGACAAACAGCGCATTTCCCTATATGACG 4062  
QY 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312  
DB 4063 CACCTTTCAGTACAGCCACCCACCAACAGCAGTTG 4098

## RESULT 4

ADH69466

ID ADH69466 standard; DNA; 4481 BP.

XX AC ADH69466;

XX 25-MAR-2004 (first entry)

XX Human spinocerebellar ataxia (SCA2) gene.

XX Spinocerebellar ataxia 2; SCA2; genetic analysis; gene; ds; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FH CDS 163..4101

FT /\*tag= a

FT /product= "Human SCA2 protein"

XX US6623927-B1.

XX 23-SEP-2003.

XX 08-NOV-2000; 2000US-00707919.

XX 08-NOV-2000; 2000US-00707919.

XX (COUL ) COUNCIL SCI &amp; IND RES.

XX Brahmachari SK, Choudhry S, Mukerji M, Jain S;

XX WPI; 2003-895465/82.

XX P-PSDB; ADH69465.

XX New oligonucleotide primers useful in predicting an individual's

XX susceptibility to spinocerebellar ataxia 2 (SCA2) disease, in genetic

XX analysis of SCA2 gene in a population, and in detecting allelic variants

XX of the SCA2 gene.

XX Disclosure; SEQ ID NO 20; 30pp; English.

XX

CC The invention relates to an oligonucleotide primer. The oligonucleotide  
CC primers are useful in molecular diagnosis, in predicting an individual's  
CC susceptibility to spinocerebellar ataxia 2 (SCA2) disease, in genetic  
CC analysis of SCA2 gene in a population and in detecting allelic variants  
CC of the SCA2 gene. The present sequence is human SCA2 gene.

XX  
SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 5,27e-200 Length: 4481

Score: 6961.00 Matches: 1312

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-10-802-228-2 (1-1312) x ADH69466 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20

DB 163 ATGCGCTCAGCGCGCGAGCTCTCGAGTCCGCGGTGGCCACCGAGTCTGCCCGCTTC 222

QY 21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArgSerGlyArg 40

DB 223 GCGGCGAGCAGGTGGCGGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCG 282

QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60

DB 283 GCGCGCGGTGGCGCGCGCGCGAGCGTATCCCTCCGCGCGCGCTCCCGCGCGCG 342

QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80

DB 343 GCGCGCGCTCCCTCCCGCGAGAGCTCGCTCCCTCCCGCTCAGACTGTTTTGGTAGAAC 402

QY 81 GlyAsnGlyGlyAlaAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyPro 100

DB 403 GGCACGCGCGCGCGCGGTTCGGCCCGCTCCCGCGCGCTCCTTGTCTCGCGCGGCT 462

QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProAlaAla 120

DB 463 CCGCGCGCTTCGTCTGCTCTCTTCCTCCCGCTCGCCAGCGCGCGCGCGCGCGCG 522

QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140

DB 523 CCAACCGCGCTCCCGCTCGCGCGCGCTCGCTCCCGCGCGCTTCCCGCGCTCTCCTTG 582

QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160

DB 583 GCGCGCGCGCTCCCGCTTCCTCCCGCGCGGTGGAGCGCGGTGTATGGGCGCTCAC 642

QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180

DB 643 ATGTCGTGAAGCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702

QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762

QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerSerSer 220

DB 763 AGCGGCTTCTAGCGTCCCGCGCGCGCGCGCTTCCCGCGCTCTCGCTCTCGCTCT 822

QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240

DB 823 TCCTCGCGCACCGCTCCCTCTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 882

QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260

DB 883 CTGGCGAGGTCGAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAACAGTAAAC 942

QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280

DB 943 ATCTATGCAAAATATGAGGATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGA 1002

QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
DB 1003 GTACAAGTGAATAATCGAGGTATATATCAAGGAGTTTTTAAAACTTACAGTCCGAAGTGT 1062  
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320  
DB 1063 GATTTGGTACTTGTATGCCGCAATCAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGT 1122  
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
DB 1123 GAAGAAATAAGGAGAGATTTTGTTCAAATGTTCCAGCTTTGTGTGTCAGTTTAAA 1182  
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
DB 1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAA 1242  
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380  
DB 1243 GTGAATGGCGAACACAAAGAGAGAGACCTGGAGCCCTGGGATGCGAGGTGAACTCACGCC 1302  
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
DB 1303 AATGAGGAACCTTGAGGCTTTGGAAATGACGTATCTAATGATGGATCCCAATGATATG 1362  
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
DB 1363 TTTTCGATATAATGAAGAAAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
DB 1423 TATACAGTGCCTTTAGAAAGAGATAACTCAGAAGAAATTTTAAAAACGGGAAAGCAAGGCA 1482  
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
DB 1483 AACCGTTAGCAGAGAATAATGAGTCAAGTGCCCAAGTACAAAGCTCGAGTGGCCCTGGAA 1542  
QY 461 AsnAspArgArgSerGluGluLubysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
DB 1543 AATGATGATAGCAGTGAGGAGAGAAAAATACACAGCAGTTTCAGAGAAATTCACGTGAACGT 1602  
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
DB 1603 GAGGGGCACAGCATAAACACTAGGGAAAAATAAATATATTCCTCTCGSACAAAGAAATAGA 1662  
QY 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
DB 1663 GAAGTCATATCTTGGGGAAGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCTGGATCG 1722  
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATTCGTGTTCA 1782  
QY 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560  
DB 1783 GACCAAGAGTAGTTAAATGGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCCTTCCTCT 1842  
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
DB 1843 CGCCACCTTCTCGTACCAGTCAGTCCCAACTCTCTTCCACTCGGGGAGCCACCCCT 1902  
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
DB 1903 ACAGCGCGCCCTCCAGGCCCTCCGCGGCGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
DB 1963 CATGGTCTCCAGCTCCTCTCTACTATGCTTAAACGCATGTCTTTCAGAAAGGGCCCTCA 2022  
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
DB 2023 AGGATGTCCCCAAGGCCCGCCGACATCTCGAANAATCACAGATTTCTGCTGGGAGGGGT 2082

QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
DB 2083 TCCATATCCAGTGGCTAGAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCTCT 2142  
QY 661 ProValAlaAaGThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680  
DB 2143 CCAGTAGCAAGGACAGTCCCTCGGGGGAACGTGGTTCATCAGTGGTTCAGTGGGTTTCCA 2202  
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
DB 2203 AGATTATCCCTTAAACTCATAGACCAGGTCTCCAGACAGAAACAGTATTTGGAAATATCC 2262  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
DB 2263 CCCAGTGGCCAGTCTTGTCTCTCCCAAGCTGATATTATTCCAACTGAAGCTGTTCGCC 2322  
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
DB 2323 ATGGCTATTCCAGCTGCAATCTCCACGCTCTAGTCTGTCATCGAACAGAGCTGTATACC 2382  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
DB 2383 CTTTCTAGTAGGCTTAAAGATTCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGG 2442  
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
DB 2443 AATAAAGAAAAATATTAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502  
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
DB 2503 GGTATATCACAGTGTGTTTCTGAAACATAGAAAAACAGATTGATGATTTAAAGAAATTTTAAAG 2562  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
DB 2563 AATGATTTTAGGTTACAGCCAAAGTTCTACTTCTGAATCTATGGATCAACTACTATAACAAA 2622  
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
DB 2623 AATAGAGAGGGAGAAAAAATCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTCTTAAG 2682  
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
DB 2683 GATTCTTTTCATTTGAAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGCAATAGC 2742  
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
DB 2743 CCCAGCATTTCCCTTCAATACTTGTAGTAACCGGACACAGAGGGGACCTGAGGTCACT 2802  
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900  
DB 2803 TCCCAAGGGGTTACAGCTTCCAGCCAGCATGTAAACAAGAGAAAGACGATAAGGAAGAG 2862  
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
DB 2863 AAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAAATCCCAATGCAAGAGAGTTTC 2922  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
DB 2923 AACCCACGTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCTCGGCTCAA 2982  
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
DB 2983 GCACAAACCTAGCCCATCTATGGTGGTTCATCAACAGGCCCACTCCAGTTTATACTCAGCCT 3042  
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
DB 3043 GTTTGTTTTCACCAAAATATGATGTATCCAGTCCCGAGTGAGCCCGGCGTCAACCTTTA 3102  
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
DB 3103 TACCCATACCTATGACGCCCATGCCAGTGAATCAAGCCAGACATATAGAGCAGTACCA 3162  
QY 1001 AsnMetProGlnGlnArgAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020

Db 3163 AATATGCCCAACAGCGGCAAGCAGCATCATCAGATGCCATGATCACCAGCGTCA 3222  
 QY 1021 AlaaAGlyProPheAlaAlaThrProAlaTyrSerThrGlnTyrValAlaTyr 1040  
 Db 3223 GCAGCGGGCCCAACGATTCAGCCACCCACAGCTTACTCCACGCAATATGTCCTAC 3282  
 QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
 Db 3283 AGTCCTCAGCAGTTCACAAATCAGCCCTTGTTTCAGCATGTGCCACATTCAGTCTCAG 3342  
 QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetAlaProProThr 1080  
 Db 3343 CATCTCTATGTCATAGTCTGTAATACAGGGAATGCTAGATGATGACACCAACA 3402  
 QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
 Db 3403 CACGCCAGCCTGGTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462  
 QY 1101 HisAlaMetTyrAlaCysProIysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
 Db 3463 CATGGCATGTATGTCATGTCCTCCAAATPACCATACACAGGAGACAGCCCTTCTTCTAC 3522  
 QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
 Db 3523 TTTGGCATTTCCAGCGGGCTCCCTTGCTCAGCAGTATGCGCACCTAACGCTACCCCTGCAC 3582  
 QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160  
 Db 3583 CCACATCTCCACACCTCAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCAACT 3642  
 QY 1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla 1180  
 Db 3643 GGTGGAAGTCATCTCCACCCAGCTCTGTTTCAGCACCATCAGCACCGCGCCAGGCT 3702  
 QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
 Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCCATTTACCCAGCGGGCTTGGCCAACT 3762  
 QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
 Db 3763 CCACCTCTCATGACACTGCTGCTCCAAACAGCAGTCCGCCACAGATAGTTTCCAGCAGCA 3822  
 QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
 Db 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGCTTCAGCGCGCGGTATACCAACCCACCC 3882  
 QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
 Db 3883 CACATGGCCCACTAGCTCAGGCTCATGTACAGTCAGGAATGTTCTTCTCATCCAACT 3942  
 QY 1261 AlaHisAlaProMetMetLeuMetThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
 Db 3943 GCCCATGCGCAATGATGCTAATGAGACACACCCCGCGGCTCCCGAGCGGCCCTC 4002  
 QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
 Db 4003 GCTCAAGTGCATACAGCCATTCAGTCTCAGCAACAGCGCATTTCCCTATATGACG 4062  
 QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312  
 Db 4063 CACCTTTCAGTACAGCCACCCACCAACAGCAGTTG 4098

## RESULT 5

ABV75291

ID ABV75291 standard; DNA; 4481 BP.

XX AC

XX AC

XX AC

DT 07-MAR-2003 (first entry)

DE Human SCA2 polypeptide encoding DNA.

XX

SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity; memory impairment; mutant; human; gene; ds.

Homo sapiens.

Key Location/Qualifiers  
CDS 163..4101  
/\*tag= a  
/product= "SCA2"

WO200289567-A2.

14-NOV-2002.

07-MAY-2002; 2002WO-US014698.

07-MAY-2001; 2001US-0289231P.

(CEDA-) CEDARS SINAI MEDICAL CENT.

Pulst SM;

WPI; 2003-103467/09.

P-PSDB; ABB82698.

New mutant mouse, useful as a disease model for identifying agents for treating obesity or memory impairment, comprises a disrupted spinocerebellar ataxia 2 gene.

Disclosure; Page 65-72; 76pp; English.

The invention relates to a mutant non-human mammal having a disrupted spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a disease model for identifying agents or compounds that may treat obesity or memory impairment. The stem cell is used for generating the mutant animal. The present sequence represents the nucleotide sequence (GenBank Accession No. U70323) encoding the human SCA2 polypeptide

Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 5,27e-200 Length: 4481  
 Score: 6961.00 Matches: 1312  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-802-228-2 (1-1312) x ABV75291 (1-4481)

QY 1 MetArgSerAlaAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
 Db 163 ATGCGCTCAGCGCGCCAGCTCTCGAGTCCGCGTGGCCACCGAGTCTCCCGCTTC 222  
 QY 21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArgSerGlyArg 40  
 Db 223 GCCGACGACAGTGGCGCGGTGGCGTCCGTCAGCGCGCGCGCGCGAGCGGCGG 282  
 QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60  
 Db 283 GCGCGCGTGGCGCGCGCGCGCGCGTATCCCTCCCGCGCGCTCCCGCGCGCGCC 342  
 QY 61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80  
 Db 343 GGGCCCCCTCCCTCCCGCGCAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGTAGCAAC 402  
 QY 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100  
 Db 403 GGCACGCGCGCGCGCGGTTCGGCCCGGCTCCCGCGGCTCTTGGTCTCGCGCGGCT 462  
 QY 101 ProArgProPheValValValLeuLeuLeuProLeuAlaSerProGlyAlaProAlaAla 120  
 Db 463 CCGCGCGCGCTTCGTCGTCGTCCTTCTCCCTCCCGCAGCGCGCGCGCTCCCGCGCGG 522

121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140  
122 |||||  
523 CCAACCGCGCTCCCGCTCGGCGCGGTGCGTCCCGCGCGGTTCGCGGTCTCCTTG 582  
524 |||||  
141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
142 |||||  
593 GCGCGCCCGGTCTCCCGGTCTCCCGCGCGGTGCGAGCGGTGTATGGGCCCCCTACCC 642  
594 |||||  
161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
162 |||||  
643 ATGTGCGTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702  
644 |||||  
181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
182 |||||  
703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762  
704 |||||  
201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerSerSer 220  
202 |||||  
763 AGCGGCCCTTCTAGCGCTCGCGCGCGCGCGCTTCGCGCTCTCGTCTCGGTCTCCTCG 822  
764 |||||  
221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240  
222 |||||  
823 TCCTGGCCACGGCTCCCTCCTCGTGGTTCGGGGGACCTCCGGCGGGGAGGCCCGCGC 882  
824 |||||  
241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
261 |||||  
883 CTGGCAGAGGTGCAAAACAGTAACAAGACTGCCTCAGTCTACGATTTCTTTTATGGA 942  
884 |||||  
261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
281 |||||  
943 ATCTATGCAAAATATCAGATGGTTTCATATCTATCATCAGTTGTTGGCTCCAAATGTGAA 1002  
944 |||||  
281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
301 |||||  
1003 GTACAGTGAAATATGGAGGTATATATGAGGAGTTTTTAAACTTACAGTCCGAGTGT 1062  
1004 |||||  
301 AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg 320  
321 |||||  
1063 GATTTGGTACTTGTATGCGGCACATGAGAAAGTAGTACAGAAATCCAGTTCGGGGCCGAAACGT 1122  
1064 |||||  
321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
341 |||||  
1123 GAAGAAATATATGGAGGTATTTTGTTCAAATGTTTCAGACTTTGTTGGTACAGTTTAAA 1182  
1124 |||||  
341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
361 |||||  
1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAA 1242  
1184 |||||  
361 ValAsnGlyGluHisLysGlyLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380  
381 |||||  
1243 GTGAATGGCGAAACAAAGAGAAGGACCTGGAGCCCTGGAGTGCAGGTGAACTCACAGCC 1302  
1244 |||||  
381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMet 400  
401 |||||  
1303 AATGAGAACTTGGAGCTTTGGAATATGACGTATCTAATGGATGGGATCCCAATCATATG 1362  
1304 |||||  
401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
421 |||||  
1363 TTTTCATATATGAAGAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
1364 |||||  
421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
441 |||||  
1423 TATACAGTGCCCTTAGAAGAGATAACTCAGAAGAAATTTTAAACGGGAAGCAAGGCA 1482  
1424 |||||  
441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
461 |||||  
1483 AACCAAGTTAGCAGAGAATAATGAGTCAAGTGCCCGAGTACAAAGCTCGAGTGGCCCTGGAA 1542  
1484 |||||  
461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
481 |||||  
1543 AATGATGATAGGATGAGGAAGAAAATACACAGCAGTTTCAGAGAAATTTCCAGTGAACGT 1602  
1544 |||||

QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
482 |||||  
Db 1603 GAGGGGCACACATAAAACACTAGGAAAAATAAATATATTCCTCTGGGCAAGAAATAGA 1662  
1604 |||||  
QY 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
521 |||||  
Db 1663 GAAGTCATATCTCTGGGAGTGGGAGACAGAAATTCACCGGTATGGGCGACGCTGGATCG 1722  
1664 |||||  
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
541 |||||  
Db 1723 GGTCTCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATTCGTGTTCA 1782  
1784 |||||  
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560  
561 |||||  
Db 1783 GACCAAAAGATAGTAAATGGAGGTGTTCCCTGGGCATCGCTTCGCCATCTCTCTTCCTCT 1842  
1843 |||||  
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
581 |||||  
Db 1843 CGCCACCTTCTCGCTACAGTCCAGTCCCAACTCTCTTCCACTTCGGGCGACCCACCT 1902  
1903 |||||  
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
601 |||||  
Db 1903 ACACGGCGCCCTCCAGGCCCTCCGGCGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
1963 |||||  
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGlyGlyProPro 620  
621 |||||  
Db 1963 CATGGTCTCCAGCTCTCTCTACTATGCCTAAACGATGTCTTTCAGAAAGGGCTCCA 2022  
2023 |||||  
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
641 |||||  
Db 2023 AGGATGTCCTCCAAAGGCCCGCAGACATCCTCGAAATCACAGATTTCTGTGGGAGGGT 2082  
2083 |||||  
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
661 |||||  
Db 2083 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCCT 2142  
2143 |||||  
QY 661 ProValAlaArgThrSerProSerSerGlyGlyThrTrpSerSerValValSerGlyValPro 680  
681 |||||  
Db 2143 CCAGTAGCAAGACAGTCCCTCGGGGGAACGTGTTCATCAGTGTGCTGGGTTCCTCA 2202  
2203 |||||  
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
701 |||||  
Db 2203 AGATTATCCCTTAAACTCATAGACCCAGGTCTCCAGACAGAACAGTATTGGAAATACC 2262  
2263 |||||  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
721 |||||  
Db 2263 CCCAGTGGGCCAGTCTTGTGCTTCTCCCCAAGCTGTGTTATTTCCAACTGAAGCTGTGCC 2322  
2323 |||||  
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
741 |||||  
Db 2323 ATGCTATTTCCAGCTGCATCTCTACGCTGTGTAGTCTGTGATCGATCGAACAGAGCTTTACC 2382  
2383 |||||  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
761 |||||  
Db 2383 CTTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATTCAGAGCAGAACTCTCTCTCGAGG 2442  
2443 |||||  
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
781 |||||  
Db 2443 AATAAGAAATATTTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502  
2503 |||||  
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
801 |||||  
Db 2503 GGTATATCACCGTGTGTTCTCGAAATAGAAACAGATTGATTTAAAGAAATTTTAAG 2562  
2563 |||||  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuAsnLys 820  
821 |||||  
Db 2563 AATGATTTTAGGTACAGCCCAAGTTCTACTTCTGAAATCTATGGATCAACTACTAACCAAA 2622  
2623 |||||  
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
841 |||||  
Db 2623 AATAGAGAGGAGAAAAATCAAGAGATTGTATCAAAAGACAAAAATTTGAACCAAGTCTTAAG 2682  
2683 |||||  
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
861 |||||



```
Db 2683 GATTCCTTTTCAATGAATAAGACAGCAACTGTACAGTGGCAGCAGCAAGCCGAAATAGC 2742
Qy 861 ProSerIleSerProSerIleLeuSerAenThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCAGCATTTCCCTTCATATCTAGTAACAGGAGCACAAGGGGACCTGAGTCACT 2802
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
Db 2803 TCCCAAGGGGTTTCAGACTTCAGCCCGCAGCATGTAAACAGAGAAAGACGATTAAGGAAGAG 2862
Qy 901 LysLysAspAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAGAGCGCAGCTGAGCAAGTTAGGAATCAACATTAATCCCAATCAAGAGAGTTTC 2922
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCACGTTCTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCTCAA 2982
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2983 GCACAACCTAGCCCATCTATGTGTGGTGCATCAACAGCCCACTCCAGTTTATATCTCAGCCT 3042
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTTGTTTTCACCAATATGATGTATCCAGTCCCGAGTCCCGAGCGCGGTGCAACCTTTA 3102
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCCCAATACCTATGAGCGCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCAACAGCGGCAAGCAGCAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGCGGGGCCACCGATGTCAGCCACCCACCGAGCTTACTCCACGCAATATGTTGCTTAC 3282
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTTTCCAAATCAGCCCTTGTTCAGCATGTGCCACATATCATGCTTCAG 3342
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCCTCATGCTATAGTCTGTAATACAGGGTAATGCTAGATGATGGCACCACCAACA 3402
Qy 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db 3403 CAGCGCCAGCCTGGTTTGTATCTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGAGC 3462
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGCATGATGATGATGCTCCAAATTTACCATACACAGAGACAAGCCCTCTTTCTTAC 3522
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGCCATTTCCAGGGGCTCCCTTGCTCAGCAGTATGGCAGCCCTTAACGCTACCTGCAC 3582
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
Db 3583 CCACATATCCACACCCCTCAGCCTTTCAGTACCCCACTGGCAGCAGCAGCAAGCAACAT 3642
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla 1180
Db 3643 GGTGGAAAGTATCTCTGACCCAGTCTCTTCAGCACCATCAGCACCAGCGGCCCGCCAGCT 3702
Qy 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTACGCCATTATACCCAGCGGGGCTTGGCCAACT 3762
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
```

---

```
Db 3763 CCACCTCCATGACACACCTGCCTCCAAACAGCAGTCCGACAGAAATAGTTTCCAGCAGCA 3822
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3823 CAACAGACTGTCTTTTACGATCCATCTCTCAGTTCAGCGCGGTATACCAACCCACCC 3882
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGGGCCACGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCAACT 3942
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGGCCCAATGCTAATGACGACACAGCCACCCGCGGTCCCCAGCGCCGCTC 4002
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db 4003 GCTCAAGTGCACTACAGCCCATTCAGTCTCGACAAGCAGCGCATTTCCCTTATATGACG 4062
Qy 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4063 CACCTTCAGTACAAGCCACCCACCAAGCAGTTG 4098
```

RESULT 6  
ADZ49164  
ID ADZ49164 standard; DNA; 4481 BP.  
XX AC ADZ49164;  
XX 30-JUN-2005 (first entry)  
DT Insulin signaling pathway related gene, SEQ ID 493.  
DE Antidiabetic; Gene Therapy; Non-insulin dependent diabetes;  
KW insulin resistance; gene; ds.  
XX Homo sapiens.  
XX US2005085436-A1.  
PN 21-APR-2005.  
XX 08-JUL-2004; 2004US-00887553.  
PF 08-JUL-2003; 2003US-0485883P.  
XX (LIHH/) LI H.  
PA (MAJU/) MA J.  
XX Li H, Ma J;  
PI WPI; 2005-305194/31.  
XX  
DR Treating, preventing or ameliorating pathological conditions associated with dysregulation of the insulin signaling pathway (ISP) comprises administering to a subject an amount of a modulator of any of the proteins regulated by ISP.  
XX Disclosure; SEQ ID NO 493; 70pp; English.  
XX The present invention relates to a method for treating, preventing or ameliorating pathological conditions associated with dysregulation of the insulin signaling pathway (ISP). The method comprises administering to a subject a modulator for ISP-regulated proteins or a pharmaceutical composition comprising the described modulator. The method is useful for treating, preventing or ameliorating pathological conditions associated with dysregulation of the ISP such as Type II diabetes or Type A syndrome of insulin resistance. The present sequence is a human homolog of a Drosophila gene regulated by ISP. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html?DocID=20050085436](http://seqdata.uspto.gov/sequence.html?DocID=20050085436).  
XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other; SQ

Alignment Scores:

Pred. No.: 5,276-200 Length: 4481  
Score: 6961.00 Matches: 1312  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-802-228-2 (1-1312) x ADZ49164 (1-4481)

QY 1 MetArgSerAlaAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
DB 163 ATGCCTCAGCGCGCGAGCTCTCGAGTCCCGGTGGCCACCGAGTCTCGCGCTTC 222  
QY 21 AlaAlaAlaArgTTPProGlyTTPArgSerLeuGlnArgProAlaArgArgSerGlyArg 40  
DB 223 GCCGAGCCAGGTGGCCCGGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGGCGG 282  
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60  
DB 283 GCGCGCGTGGCGCGCGCGCGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCC 342  
QY 61 GlyProProProSerArgGlnSerSerProProProSerAlaSerAspCysPheGlySerAn 80  
DB 343 GCGCGCGCTCCCGCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTTGGTAGCAAC 402  
QY 81 GlyAenGlyGlyAlaAlaPheArgProGlySerArgAlaLeuLeuGlyGlyPro 100  
DB 403 GCGACGCGCGCGCGGTTCGCGCGCGCTCCCGCGCGCTCTCTGGTCTCGCGGGCT 462  
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120  
DB 463 CCGCGCGCTTCGTCTGTCTCTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCG 522  
QY 121 ProThrArgAlaSerProLeuGlyAlaAlaArgAlaSerProProArgSerGlyValSerLeu 140  
DB 523 CCAACCGCGCGCTCCCGCTCGCGCGCGTGGTCCCGCGCGCTCCCGCGCTCTCTCTG 582  
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
DB 583 GCGCGCGCGCTCCCGCTGTCCCGCGCGCGCTCGCGCGCTGTATGGCGCGCTCACC 642  
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
DB 643 ATGCTGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702  
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762  
QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220  
DB 763 AGCGCGCTTCTAGCGTGGCGCGCGCGCGCGCTTCGCGCTCTCGTCTCTCGTCTCTCG 822  
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240  
DB 823 TCCTCGCGCCAGCGCTCCCTCTCGTGTGGCGCGACCTCCGCGCGCGCGCGCGCGCG 882  
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
DB 883 CTGGCAGAGTGCAGAAACAGTAACAAAGGACTGCGCTCAGTCTACGATTTCTTTTATGGA 942  
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
DB 943 ATCTATGCANATATGAGGAATGTTCTATATCTTACTTACATCAGTTGTGGTCTCCAAATGTGAA 1002  
QY 281 ValGlnValLysAenGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
DB 1003 GTACAAGTGAATAATGGAGGTATATATGAAGGAGTTTTTAAAACTTACAGTCCGAAGTGT 1062  
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320

DB 1063 GATTTGGTACTTGATGCCGCACATGAGAAAGTAGTACAGAAATCCAGTTTCGGGGCCCCAAACGT 1122  
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys 340  
DB 1123 GAACAATAATGGAGAGTATTTTCTCAAAATGTTTCAGACTTTGTGTGGTACAGTTTAA 1182  
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
DB 1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA 1242  
QY 361 ValAenGlyGluHisLysGluLysAspLeuGluProTTPAspAlaGlyGluLeuThrAla 380  
DB 1243 GTGAATGGCAACAACAAGAGAGAGACTTGAGCGCTGGGATGAGGTGAACACTCACAGCC 1302  
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTTPAspProAsnAspMet 400  
DB 1303 AATCAGGACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATG 1362  
QY 401 PheArgTyrAenGluGluAenTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
DB 1363 TTTCCATATATGAAGAAAAATATGCTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
DB 1423 TATACAGTGGCTTTAGAAAGAGATAACTCAGAAAGATTTTAAACCGGAAGCAAGGGCA 1482  
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
DB 1483 AACCACTTAGCAGAAGAAATTTGAGTCAAGTGCAGTCAAAAGCTCGAGTGGCCCTGGA 1542  
QY 461 AsnAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
DB 1543 AATGATGATAGCAGTGAAGAAAAATACACAGCAGTTACAGAAAAATTCAGTGAACGT 1602  
QY 481 GluGlyHisSerIleAenThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
DB 1603 GAGGGGCACAGCATAAACACACTAGGGAATAATAATATATCTCTCTGGAACAAAGAAATAGA 1662  
QY 501 GluValIleSerTTPGlySerGlyArgGluAsnSerProArgMetGlyGlnProGlySer 520  
DB 1663 GAAGTCATATCTCGGGAGTGGGAGACAGAAATTCACCGGTATGGGCGAGCTGGATCG 1722  
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTCAACCCCGAATTCCTGGTTCA 1782  
QY 541 AspGlnArgValValAsnGlyGlyValProTTPProSerProCysProSerProSerSer 560  
DB 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCCATCGCCTTGCCCATCTCTCTCTCT 1842  
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
DB 1843 CGCCACCTTCTCGCTACCGTCCAGTCCAGTCCCAACTCTCTCCACCTCGGCGACCCACCT 1902  
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
DB 1903 ACACGGCGCGCTCCAGCGCGCGCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
DB 1963 CATGGTCTCCAGTCTCTGCTCTACTATGCTTAACGCATGTCTTTCAGAAAGGCGCTCCA 2022  
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
DB 2023 AGGATGTCCTCCAAAGGCCCGACATCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT 2082  
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
DB 2083 TCCATATCCAGTGGCTTAGAATTTGTATCCCAACCCAGCCAGTGAAGCAGTACTCTCT 2142  
QY 661 ProValAlaArgThrSerProSerGlyThrThrProSerSerValSerGlyValPro 680  
DB 2143 CCAGTAGCAAGGACAGTCCCTCGGGGGGAACGTTGGTTCATCAGTGGTTCAGTGGGGTTCCA 2202

```
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTATCCCTAAAACTCATAGACCAGGCTCTCCAGACAGAAAGTAGTATGGAATATCC 2262
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720
Db 2263 CCAGTGGGCCAGTTCTTGCTTCTCCCAAGCTGGTATTTATTCCTCAACTGAACTGTGGCC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Db 2323 ATGCCTATTCCAGCTGCATCTCTCCAGCTGCTAGTCTGCTGCATCGAACAGAGCTGTATCC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Db 2383 CCTTCTAGTGGGCTAAAGATTCCAGGCTTCAAGATCAGAGGAGCAAACTCTCTCGAGGG 2442
QY 761 AsnLysGluAsnLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db 2443 AATAAAGAAATATTAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
Db 2503 GGTATATCACCACTGTTCTGACATAGAAACAGATTGATGATTTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db 2563 AATGATTTTAGGTTACAGCCAAAGTTCTACTTCTGAATCTATGATCAACTACTATAACAAA 2622
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
Db 2623 AATAGAGAGGGAGAAATTAAGAGATTGATCAAGACAAAATTTGAACAAGTGTCTAAG 2682
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db 2683 GATTCCTTCATTGAATAATAGCAGCACTGTATCCAGTGGCAGCAGCAAGCCGATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCCAGCATTTCCCTCTCAATCTAGTTAGTAAACCGAGCAGCAAGAGGGGACCTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
Db 2803 TCCCAAGGGGTTCCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGATAAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAGACGCAGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCAGTTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2983 GCACAACTAGCCCATCTATGGTGGGTCATCAACAGCCAACTCCAGTTTATCTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTTGTTTGACCAAAATATGATGATATCCAGTCCAGTGAGCCAGGCGTGCAACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCAATACCTATGAGCCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCAACAGCGCGCAAGACCAACATCATCAGAGTGCCCATGATGCAACCCAGCGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGCGGGCCACCGATTGACGCCACCCACAGCTTACTCCAGCAATATGTTGCTTAC 3282
```

```
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCCTCATGTCTATAGTCTGTAAATACAGGTAATGTAGTAATGATGGCACCACCAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyValHisGluGlnThr 1100
Db 3403 CACGCCCCAGCGTGGTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGGCATGTATGCATGTCCCAAAATACATACACAGGAGCAAGCCCTCTTCTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGGCAATTTCCACGGGCTCCCTTGTCTCAGCAGTATGGCACCTTAACGCTACCTGTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 3583 CCACATACTCCACACCCCTCAGCTACCCCTCAGTACCCCTGAGCAGCAGCAAGCCCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Db 3643 GGTGAAGTCACTCTTCACCCAGTCTGTTCAGCACCATTACAGCAGCCGCGCCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTTACCACGGGGCTTGCCCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3763 CCACCTCCTCAGCAGCCTGCTCCAAACCGCAGTCGCCACAGAAATAGTTTCCCGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3823 CAAACAGACTGTCTTTACGATCCATCTCTCAGCTTACCGCGCGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGGCCCCGCTACCTCAGGCTCATGTACAGTCAGGAATGGTTCTCTTCTCATCCAAC 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGCGCCAAATGATGCTAATGACGACACAGCCACCCCGCGTCCCCAGCGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db 4003 GCTCAAGTGCATCAGGCCCATTCAGTCTCGAACACAGCGCATTTCCCTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4063 CACCCTTCAGTACAGCCCAACCCACCAACAGCAGTTG 4098
RESULT 7
AAV30270
ID AAV30270 standard; DNA; 4367 BP.
XX
AC AAV30270;
XX
XX 02-OCT-1998 (first entry)
DT
XX
DE Gene causative of spinocerebellar ataxia type 2 (SCA2) DNA sequence.
XX
XX Spinocerebellar ataxia type 2; SCA2; gene therapy; antisense therapy;
KW CAG repeat; neurodegenerative disease; ds.
XX
OS Homo sapiens.
XX
XX Location/Qualifiers
PH 49..3990
FT CDS
```

```

FT      /*tag= a
FT      /product= "Spinocerebellar ataxia type 2 associated
FT      protein"
FT      repeat_region 544..612
FT      /*tag= b
FT      /*note= "normal CAG repeat region; this is increased in
FT      patients with SCA2"
FT      repeat_unit 544..546
FT      /*tag= c
FT      XX
FT      PN W09818920-A1.
FT      XX
FT      PD 07-MAY-1998.
FT      XX
FT      PF 30-OCT-1997; 97WO-JP003946.
FT      XX
FT      PR 30-OCT-1996; 96JP-00304059.
FT      XX
FT      PA (SRLS-) SRL INC.
FT      XX
FT      PI Tsuji S, Sanpei K;
FT      XX
FT      DR WPI; 1998-272215/24.
FT      XX
FT      DR P-PSDB; AAW60213.
FT      XX
FT      XX Nucleic acid fragments associated with spinocerebellar ataxia type 2 -
FT      contain increased number of CAG repeat region compared to normal gene.
FT      XX
FT      PS Claim 1; Page 13-22; 38pp; Japanese.
FT      XX
FT      CC This represents the sequence of a gene causative of spinocerebellar
FT      CC ataxia type 2 (SCA2), a neurodegenerative disease. This gene associated
FT      CC with SCA2, has a tri-nucleotide (CAG) repeat region which in the
FT      CC expression product produces a polyglutamine sequence from Gln-166 to Gln-
FT      CC 188. In the normal gene there are 15-25 CAG repeats but in SCA2 patients
FT      CC this number is increased to 35-100. Peptides encoded by nucleic acid
FT      CC fragments (DNA or RNA) containing sequences from the SCA2 associated
FT      CC gene, antibodies recognising the peptides and antisense nucleic acids
FT      CC hybridising with the nucleic acid fragments can be used for the
FT      CC investigation and diagnosis of SCA2. They can also be used for the
FT      CC treatment of SCA2 by antisense therapy or gene therapy
FT      XX
FT      SQ Sequence 4367 BP; 1124 A; 1328 C; 991 G; 924 T; 0 U; 0 Other;
FT      XX

Alignment Scores:
Pred. No.: 1,066-199 Length: 4367
Score: 6950.50 Matches: 1312
Percent Similarity: 99.92% Conservative: 0
Best Local Similarity: 99.92% Mismatches: 0
Query Match: 99.85% Indels: 1
DB: 2 Gaps: 1

US-10-802-228-2 (1-1312) x AAV30270 (1-4367)

QY      1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB      49 ATGCCTCAGCGCGCGAGCTCTCGAGTCCCGGTCGCGGTCACCGAGTCTCGCGCTTC 108
QY      21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg 40
DB      109 GCCGAGCAGGTGCCCGGGGGCGCTCGCTCCAGCGCGCGCGCGCGAGCGGGCGG 168
QY      41 GlyGlyGlyValAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB      169 GSCGCGGTGGCGCGCGCGCGGACCGTATCCCTCGCGCGCCCTCCCGCGCGCGCGCC 228
QY      61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB      229 GSCCGCGCTCCCGCGCAGAGTCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 288
QY      81 GlyAsnGlyGlyValAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyPro 100
DB      289 GGCACGCGCGCGCGCGCTTCGCGCGCGCTCCCGCGCGCTCTCTTGGTCTCGCGCGGCT 348

```

```

QY      101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB      349 CCGCGCGCTTTCGTCTGTCTCTTCTCCCTCGCAGACCCCGGGCGCCCTCCGCGCGCG 408
QY      121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB      409 CCAACCCGCGCTCCCGCTCGCGCGCTCGCTCCCGCGCGCTTCGCGCGCTCTCTCTG 468
QY      141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB      469 GCGCGCGCGCTCCCGGCTGTCCCGCGCGCGGTGCGAGCGGTGTATGGGCCCCCTCAC 528
QY      161 MetSerLeuLysPro--GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 179
DB      529 ATGTCTGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 588
QY      180 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
DB      589 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 648
QY      200 GlySerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerValSer 219
DB      649 GGCAGCGCGCTTCTAGCTGCCCGCGCGCGCTTCGCGCTCTCTCTCTCTCTCTCT 708
QY      220 SerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgPro 239
DB      709 TCGTCTCTCGGCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 768
QY      240 GlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAsp 259
DB      769 GGCCTTGGCAGAGGTGCAAAACAGTAACAAAGAGCTGCTCAGTCTACGATTTCTTTGAT 828
QY      260 GlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCys 279
DB      829 GAATCTATGCAATATAGGAGTGTTCATATCTTACATCAGTTGTTGGTCTCCAAATGT 888
QY      280 GluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLys 299
DB      889 GAAGTACAAGTGAAAAATGGAGGTATATATCAAGAGAGTTTTTAAAACTTACAGTCCGA 948
QY      300 CysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLys 319
DB      949 TGTGATTTGGTACTTGTATGCGCGCATGAGAAAAAGTACAGAAATCCAGTTCGGGGCGG 1008
QY      320 ArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPhe 339
DB      1009 CGTGAAGAAATATGGAGAGTATTTTGTTCAAATGTTCCAGACTTTTGTGTGTACAGTTT 1068
QY      340 LysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAla 359
DB      1069 AAAGATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCT 1128
QY      360 LysValLysGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThr 379
DB      1129 AAAGTGAATGGCGAACACAAAGAGAGAGACCTGGAGCGCTCGGATGCGAGTGAACCTCA 1188
QY      380 AlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAsp 399
DB      1189 GCCAATGAGAACTTGGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGAT 1248
QY      400 MetPheArgTyrAsnGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSer 419
DB      1249 ATGTTTCGATATATGAAGAAAAATATGGTGTAGTGTCTACGTATGATAGCAGTTATCT 1308
QY      420 SerTyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArg 439
DB      1309 TCGTATACAGTGCCTTAGAAAAAGAGATAACTCAGAAAGAAATTTTAAAAACGGGAAGCA 1368
QY      440 AlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeu 459
DB      1369 GCAAACCGATTAGCAGAAAGAAATTTGAGTCAAGTCCCGAGTACAAAGCTCGAGTGCCCTG 1428

```



```

Db      3589  GCTCTCCATCTGGCAGTCCACAGCAGCAGTCAGCCATTTACACGCGGGCTTGGCCCA 3648
Qy      1200  ThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAla 1219
Db      3649  ACTCCACCCCTCATGACACCTGCTCCAAACACGCGAGTCGCCACAGAATAGTTTCCACGCA 3708
Qy      1220  AlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyThrAsnPro 1239
Db      3709  GCACAAACAGACTGTCTTTACATCATCTCTCAGGTTTCAGCCGGGTATACCAACCCCA 3768
Qy      1240  ProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisPro 1259
Db      3769  CCCACATGGCCACGTACTCAGGCTCATGTACAGTCAGGAATGTTCTTCTTCATCCA 3828
Qy      1260  ThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAla 1279
Db      3829  ACTGCCATGGCCCAATGATGCTAATGACGACACAGCCACCCGGCGGTCCCGAGCGGCC 3888
Qy      1280  LeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyMet 1299
Db      3889  CTCGCTCAAGTGTCACTACAGCCCATTCAGTCTCGACAAACAGCGCATTTCCCTATATG 3948
Qy      1300  ThrHisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db      3949  ACGCACCTTCAGTACAGCCACCCACCACACAGCAGTTG 3987

RESULT 8
AAAT78912
ID      AAAT78912 standard; cDNA; 4200 BP.
XX
AC      AAAT78912;
XX
DT      09-FEB-1998 (first entry)
XX
DE      Spinocerebellar ataxia gene SCA2.
XX
KW      Monoclonal antibody; neurodegenerative disease; polyglutamine; TBP;
KW      repeat region; affinity; TATA binding protein; Kennedy disease;
KW      transcription initiation factor; lymphoblastic cell line; schizophrenia;
KW      Huntington's disease; dominant autosomal spinocerebellar ataxia;
KW      x-linked spino-bulbar muscular atrophy; familial spastic paraplegia;
KW      dentatorubral-pallidolusial atrophy; bipolar affective disorder;
KW      manic depressive psychosis; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      misc_feature
FT      1. 3986 Location/Qualifiers
FT      /tag= g
FT      /note= "sequence contained in DAN1 clone"
FT      3. .2747
FT      /tag= a
FT      /product= "SCA2 protein"
FT      /note= "this CDS contains a putative translational start
FT      codon for the SCA2 protein at positions 243-245"
FT      3. .242
FT      /tag= c
FT      /note= "putative open reading frame which is in frame
FT      with the putative translational start site of the SCA2
FT      open reading frame"
FT      239. .245
FT      /tag= d
FT      /note= "putative Kozak consensus signal"
FT      258. .323
FT      /tag= e
FT      /note= "encodes polyglutamine repeat region; contains
FT      repeats of CAG with 2 CAA codons interspersed"
FT      258. .260
FT      /tag= f
FT      /note= "CAG repeats"
FT      2594. .3640
FT      /tag= b
FT      /note= "this second open reading frame may be derived by

```

```

FT      misc_feature
FT      3987. .4200 a frameshift or by alternative splicing"
FT      /tag= h
FT      /note= "derived from the EST's AAH92640, AAN90240 and
FT      AA213574 from dbEST database"
FT      4023. .4029
FT      misc_feature
FT      /tag= i
FT      /note= "region which differs in length between the
FT      sequences of the EST clones AAH92640, AAN90240 and
FT      AA213574"
XX
XX      WO9717445-A1.
XX
XX      15-MAY-1997.
XX
XX      08-NOV-1996; 96WO-FR001773.
XX
XX      10-NOV-1995; 95FR-00013576.
XX
XX      (CNRS ) CNRS CENT NAT RECH SCI.
XX      (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX      Tora L, Lutz Y, Trottier Y, Mandel J;
XX
XX      WPI; 1997-281034/25.
XX      P-PSDB; AAW24800, AAW24801.
XX
XX      Antibody 1C2 used for treating or preventing neuro-degenerative diseases
XX      - associated with proteins containing long poly:glutamine repeats, e.g.
XX      Huntington's disease.
XX
XX      Claim 21; Page 45-47; 69pp; French.
XX
XX      The invention relates to a monoclonal antibody (Mab) 1C2 for the
XX      treatment of neurodegenerative diseases associated with the presence of
XX      polyglutamine repeat regions. This Mab is already known for its affinity
XX      to the TATA binding protein (TBP) transcription initiation factor,
XX      especially at the amino acid sequence LEEQQRQQQQQ found at the N-
XX      terminus of TBP. Mab 1C2 has been shown to have a high affinity for
XX      polyglutamine repeats with a proportional affinity to the number of
XX      glutamine repeats. This affinity has been used to identify genes encoding
XX      proteins containing long polyglutamine repeats which are implicated in
XX      neurodegenerative diseases. A screen of an expression library, generated
XX      from a lymphoblastic cell line from a patient suffering from
XX      spinocerebellar ataxia (SCA), with Mab 1C2 isolated 6 new sequences
XX      (AAAT78906-T78911) encoding polyglutamine repeats. Mab 1C2 also isolated
XX      the complete SCA2 gene in clone DAN1 (sequence presented here). The
XX      sequence appears to contain 2 open reading frames (ORF) the second of
XX      which may be generated by an frameshift slippage or by an alternative
XX      splicing event. The first ORF also encodes a 22 amino acid polyglutamine
XX      repeat region near the N-terminus of the protein. Normal SCA2 alleles
XX      contain 17-29 CAG triplet repeats with 1-3 CAA repeats interspersed
XX      whereas the mutant sequence from patients with SCA contains at least 30,
XX      preferably 37-50 CAG repeats. Mab 1C2, active fragment of it or nucleic
XX      acids encoding it are specifically used to treat Huntington's disease,
XX      SCA types 1-5 or 7, X-linked spino-bulbar muscular atrophy (Kennedy
XX      disease), dentatorubral-pallidolusial atrophy, dominant autosomal
XX      spinocerebellar ataxia, familial spastic paraplegia, bipolar affective
XX      disorder, manic depressive psychoses and schizophrenia
XX
XX      Sequence 4200 BP; 1152 A; 1200 C; 913 G; 935 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 1.49e-182 Length: 4200
XX      Score: 6377.00 Matches: 1211
XX      Percent Similarity: 98.30% Conservative: 0
XX      Best Local Similarity: 98.30% Mismatches: 1
XX      Query Match: 91.61% Indels: 21
XX      DB: 2 Gaps: 1
XX
XX      US-10-802-228-2 (1-1312) x AAAT78912 (1-4200)
XX
XX      Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgLeuLeuGlyGlyPro 100

```

Db 3 GGCACCGCGCGCGCGTTCGCGCCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGCGCT 62  
Qy 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120  
Db 63 CCCCGCCCTTGGTGTCTCTTCCCTCCCGCGCGCGCGCGCGCGCGCGCG 122  
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140  
Db 123 CCAACCGCGCGCTCCCGCGCGCGTGGCGTCCCGCGCGTTCGCGCGCTCTCTCTG 182  
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
Db 183 GCGCGCCGCGCTCCCGCGCTGTCGCGCGCGGTGGAGCGGTGTATGGCGCCCTCAC 242  
Qy 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
Db 243 ATGTCTGAGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 302  
Qy 181 GlnGlnGlnGlnGlnGlnProProAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 200  
Db 303 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362  
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSer 220  
Db 363 AGCGCGCTTCTAGCGTCCCGCGCGCGCGCTTCGCGCTCTCTCTCTCTCTCT 422  
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaAlaAlaAlaAlaAlaAla 240  
Db 423 TCTTGGCCACCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482  
Qy 241 LeuGlyArgGlyArgAsnSerAsnLeuGlyLeuProGlnSerThrIleSerPheAspGly 260  
Db 483 CTGGCAGAGGTGCAACAGTAAACAGGACTGCCTCAGTCTACGATTCTTTTGTATGA 542  
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
Db 543 ATCTATGCAAAATATGAGGATGTTTCTATATCTATCTATCTATCTATCTATCT 602  
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
Db 603 GTACAAAGTGAANAATGAGAGTATATATGAGAGATTTTAAACCTTACAGTCCGAGTGT 662  
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320  
Db 663 GATTTGGTACTTGTATGATCCACATGAGAAAAGTACAGAAATCCAGTTCCGGCGCGAAAG 722  
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys 340  
Db 723 GAAGAAATATGAGAGATTTTGTTCAAATGTTTACAGCTTTGTGTGTACAGTTTAA 782  
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
Db 783 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTCTCTATCAGTGTCTAA 842  
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380  
Db 843 GTGAATGGCGCAACACAAAGAGAGGACCTGGAGCCCTGGGATCGAGTGAATCTCACAGCC 902  
Qy 381 AsnGluGluLeuAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
Db 903 AATGAGAACTTGGAGCTTTGGAAATGACGTATCTTAATGGATGGATCCCAATGATATG 962  
Qy 401 PheArgTyrAsnGluLeuLeuTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
Db 963 TTTTCGATATAATGAAGAAATATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022  
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaAlaArgAla 440  
Db 1023 TATACAGTCCCTTAGAAGAGATACTCAGAGAAATTTTAAACCGGAGCAAGCGCA 1082  
Qy 441 AsnGlnLeuAlaGluLeuGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460

Db 1083 AACCAAGTTAGCAGAGAAATTTGAGTCAAGTCCCAAGTACAAAGCTCGAGTGGCCCTGGAA 1142  
Qy 461 AsnAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
Db 1143 AATGATGATAGAGTGGAGAGAAATACACAGCAGTTTCAGAGAAATTCAGTGAACCT 1202  
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyClnArgAsnArg 500  
Db 1203 GAGGGGCACAGATAAACACTAGGAAATATAATATATTCCTCCCTGGACAAAGAAATAGA 1262  
Qy 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
Db 1263 GAAGTCATATCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCG 1322  
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
Db 1323 GGTCTCATGCCATCAAGATCCACTTCTCACATTCAGATTTCAACCCGAATTTCTGGTTCA 1382  
Qy 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560  
Db 1383 GACCAAGAGTAGTTAATGGAGGTGTTCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCT 1442  
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
Db 1443 CCCCCACCTTCTCGCTACAGTCAGGTCCCAACTCTCTTCCACCTCGGCAGCCACCCCT 1502  
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
Db 1503 ACACGCGCGCTCCAGCGCCCTCCGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1562  
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Db 1563 CATGGTTCTCCAGCTCCTGTCTACTATGCTTAAACGCATGCTTTCAGAGGGCTCCA 1622  
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Db 1623 AGGATGTCCCCAAGGCCAGCAGCATCTCTCGAAATCACAGAGTTTCTCTGGGAGGGT 1682  
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
Db 1683 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGTACTCTCT 1742  
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTyrSerSerValValSerGlyValPro 680  
Db 1743 CCAGTAGCAAGACCCAGTCCCTCGGGGGAAGCTGTGTCATCAGTGTCTAGTGGGGTTCCA 1802  
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
Db 1803 AGATTATCCCTTAAACCTCATAGACCCAGTCTCCAGACAGAACAGTATTGGAAATACC 1862  
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
Db 1863 CCCAGTGGGCCAGTCTTCTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGCC 1922  
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
Db 1923 ATGCCTATTTCAGCTGTCATCTCAGCCTGTAGTCTCTCGATCGAACAGAGTGTATACC 1982  
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
Db 1983 CTTCTTAGTGAGCTTAAAGATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCTCGAGG 2042  
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
Db 2043 AATAAGAAATATTAACCCCAATGAACATCACCTAGTCTCTCAAAAGCTGAACAA 2102  
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
Db 2103 GGTATATCACAGTGTCTTCTGAACATAGAAACAGATTTGATGATTTAAAGAAATTAAG 2162  
Qy 801 AsnAppPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
Db 2163 AATGATTTTAGGTTACAGCCAAAGTTCTACTTCTGAAATCTATGATCAACTACTAAACAA 2222



```
Qy 821 AsnArgGluGlyGluLysSerArgAspLeuLeuIleLysAspLysIleGluProSerAlaLys 840
Db 2223 AATAGAGAGGGAGAAAAATCAAGAGATTGTATCAAGAGCAAAAATTGAACCAAGTGCTAAG 2282
Qy 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db 2283 GATTCTTTTCATTTGAAAAATAGCAGCAGCAACTGTACACAGTGCAGCAGCAAGCCGAATAGC 2342
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2343 CCCAGCATTTCCCTTCAATACTTAGTAACAGGAGCAAGAGGGACCTGAGGTCACT 2402
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
Db 2403 TCCCAAGGGGTTTCAGACTTCAGCCCGCAGCATGTAACAAGAGAAAGACGATAAGGAAG 2462
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2463 AAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCCAAAGGAGTTC 2522
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2523 AACCCACGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCTCAA 2582
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2583 GCACAACCTAGCCCATCTATGTGGGTTCATCAACAGGCCAACTCCAGTTTATACTCAGCCT 2642
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 2643 GTTTGTTTGTGACCAAAATATGATGTATCCAGTCCAGTGAGCCAGCGGTGCAATA---- 2698
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 2699 -----CCCA 2701
Qy 1001 AsnMetProGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 2702 AATATGCCCAACAGCGGCAAGACCGCATCATCAGAGTGCCATGATGACCCCGGTCA 2761
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 2762 GCACGGGGCCACCGATTGCGAGCACCCACAGCTTACTCCAGCAATATGTGCGCTAC 2821
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 2822 AGTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAG 2881
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 2882 CATCTCATGTCTATAGTCTGTAATACAGGGTAATGCTAGATGATGGCACCAACA 2941
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db 2942 CACGCCAGCCTGTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGC 3001
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3002 CATCGGATGTATGATGATGCCAAATACCATACAAAGGAGACAGCCCTTCTTCTTCTAC 3061
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3062 TTTGCCATTTCCAGGGGCTCCTTGTCTCAGCAGTATGCGCACCTTAAGCTTACCTGCGAC 3121
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 3122 CCACATACTCCACACCTTCAGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 3181
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180
Db 3182 GGTGGAGTCTATCTGACCCAGTCTCTGTTTCAGCACCATCATGACACCGCCGCCAGGCT 3241
```

```
Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3242 CTCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTACCACGCGGGCTTTGGCCAACT 3301
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3302 CCACCTCTCCATGACACCTGCTCCAAACACGAGTGCACAGAAATAGTTTCCAGCAGCA 3361
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3362 CAACAGACTGTCTTACGATCCATCTTCTCAGCTTACGCGGGCTATACCAACCCACCC 3421
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3422 CACATGCCCAACGTACCTCAGCTCATGTACAGTCAGAAATGGTTCTCTTCTCATCCAAC 3481
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3482 GCCCATGCGCAATGATGCTAATACACACACAGCAGCCGCGGTCCCCAGGCGGCTC 3541
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db 3542 GCTCAAGTGCATACAGCCCATTTCCAGTCTCGACACAGCGCATTTCCCTATATGAGC 3601
Qy 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312
Db 3602 CACCCTTCAGTACAAAGCCACCAACAGCAGTTG 3637
RESULT 9
ABV75290
ID ABV75290 standard; DNA; 4225 BP.
XX
AC ABV75290;
XX
DT 07-MAR-2003 (first entry)
XX
DE Mouse SCA2 polypeptide encoding DNA.
XX
KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;
XX memory impairment; mutant; mouse; gene; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
CDS 27..3884
FT /*tag= a
FT /transl_except= (pos: 222..224, aa: Xaa)
FT /note= "Xaa is Pro"
FT /product= "SCA2"
XX
WO200289567-A2.
XX
PD 14-NOV-2002.
XX
PF 07-MAY-2002; 2002WO-US014698.
XX
PR 07-MAY-2001; 2001US-0289231P.
XX
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI Pulst SM;
XX
DR WPI; 2003-103467/09.
DR P-PSDB; ABB82697.
XX
PT New mutant mouse, useful as a disease model for identifying agents for
PT treating obesity or memory impairment, comprises a disrupted
PT spinocerebellar ataxia 2 gene.
XX
PS Disclosure; Page 56-62; 76pp; English.
XX
CC The invention relates to a mutant non-human mammal having a disrupted
CC spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a
```



Db 1842 TCTGCTGGAGAGGCTCATGCTAGTGCCTAGAAATTGTATCCCAATCCCCCAAGT 1901  
Qy 656 GluAlaAlaThrProValAlaArgThrSerProSerGlyGlyThrTyrSerSerVal 675  
Db 1902 GAAGCAGCTGCTCTCCAGTGGCAAGGACCACTCTCGAGGGGGAACGTGCTCTCAGTG 1961  
Qy 676 ValSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsn 695  
Db 1962 GTCAGTGGGGTTCCAAGGTTATCTCCCAAAATCTCACAGCCAGCGTCTCCAGGCGAGAGC 2021  
Qy 696 SerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleLeuPro 715  
Db 2022 AGCATTGAATACTCTCCAGCGGGCTGTGCTGCTCTCCCAAGCTGGCATATCCCT 2081  
Qy 716 ThrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSer 735  
Db 2082 GCAGAAGCCGTTTCCATGCCTGTTCCCGCGCATCTCCGACTCTCTGCCAGCCCTGCATCC 2141  
Qy 736 AsnArgAlaValThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGln 755  
Db 2142 AACAGAGCACTGACCCCATCTATTGAGGCAAAAGATTCCAGGCTTCAAGATCAGAGGCAG 2201  
Qy 756 AsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSer 775  
Db 2202 AACTCTCTCGAGGAGTAAAGAAATGTTAAAGCAAGTGAACATCACTAGCTTTTCA 2261  
Qy 776 LysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAsp 795  
Db 2262 AAAGCTGACAAAGGTATGTACCAGTTGTTCTGAACACAGAAACAGATTGATGAC 2321  
Qy 796 LeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAsp 815  
Db 2322 TTAAAGAAAGTTTAAAGATGATTTAGGTTACAGCCAAGCTCTACATCTGAATCTATGGAT 2381  
Qy 816 GlnLeuLeuAsnLysAsnArgGluGlyGlyLysSerArgAspLeuIleLysAspLysIle 835  
Db 2382 CAACTACTTAAGCAAAATAGAGAGGAGAAAGTCACAGATTTTGATTAAGATAAAACG 2441  
Qy 836 GluProSerAlaLysAspSerPheIleGlu-----AsnSerSerSerAsnCysThrSer 853  
Db 2442 GAAGCAAGTGTAGGATAGTTTCAATTCACAGCAGCAGCAGCAGCACTGTACCAGT 2501  
Qy 854 GlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThrGluHis 873  
Db 2502 GGCAGCAGCAAGACCAACAGCCCTAGCATCTCCCTTCATGCTTAGTAATGCAGAGCAC 2561  
Qy 874 LysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCysLysGln 893  
Db 2562 AAGAGGGGCTGAGGTACATCCCAAGGGGTGCAGACTTCCAGCCCAAGCCTGCANAAACAA 2621  
Qy 894 GluLysAspAspLysGluGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeu 913  
Db 2622 GAGAAGGATGACAGAGAGAGAGAGAAAGACACACAGAGCAGGTTAGGAATTCGACATTG 2681  
Qy 914 AsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLysProSerThrThr 933  
Db 2682 AATCCCAATGCAAGAGGATCAACCCCTGTTCTTCTCTCAGGCCAAAGCCCTTCTACTACC 2741  
Qy 934 ProThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnPro 953  
Db 2742 CCAAGCTCACCTCGGCCCTCAAGACACAAACCCAGCCCATCTATGTTGGGTTCATCAGAGCCA 2801  
Qy 954 ThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal 973  
Db 2802 GCTCCAGTGTACACTCAGCCCTGTGTGCTCGCACCAATATGATGATATCCCTCCAGTG 2861  
Qy 974 SerProGlyValGlnProLeuTyrProIleProMetThrProMetProValAsnGlnAla 993  
Db 2862 AGCCCGGGCGTACAACTTTATACCAATACCTATGACGCCCATGCTGTGAACCAAGCC 2921  
Qy 994 LysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAspGlnHis 1011

Db 2922 AAGACATATAGACAGCAGGTAAGTACCATAATATGCCCCAACAGCAGCAAGACCAATCATCAT 2981  
Qy 1012 GlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThrProPro 1031  
Db 2982 CAAAGCACCATGATGACCCAGCCTCGCGCAGCGCCACCATCTGTAAGCACCACCGGCC 3041  
Qy 1032 AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGlnProLeuVal 1051  
Db 3042 GCTTACTCCATCTAGTAGTTCCTTACAGCCCTCAGCAGTTTCCCAATTCAGCCTTTGGTC 3101  
Qy 1052 GlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGly 1071  
Db 3102 CAGCATGTGCGCATATTACGTCTCAGCATCTCTCATGTGTACAGTCTGTCTATCAAGGT 3161  
Qy 1072 AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAla 1091  
Db 3162 AATGCCAGGATGATGGCACCACAGCATGTCTCAGCCTGGTTAGTGTCTCTTTCAGCT 3221  
Qy 1092 ThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyr 1111  
Db 3222 GCTCAGTTGCGGGTCTCAGAGCAGCAGCCCATGTATGCATGTCCCAATTTACCATAC 3281  
Qy 1112 AsnLysGluThrSerProSerPheAlaIleSerThrGlySerLeuAlaGlnGln 1131  
Db 3282 AACAGGAGACAGCCCTTCTTCTTGTGCAATTTCCACCGGCTCCCTCGCTCAGCAG 3341  
Qy 1132 TyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThr 1151  
Db 3342 TATGCATCCTAATGCGGCTGCATCCACATACTCCCCATCTCTCAGCCTTCGCGCAC 3401  
Qy 1152 ProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaProSerProValGln 1171  
Db 3402 CCCACCGCAGCAGCAGCAAAAGCCAGCATGTTGGAAGTCACTGTCACCGCAGTCTGTTT 3461  
Qy 1172 HisHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGlnGlnGlnSerAla 1191  
Db 3462 CACCATCAGCACCAAGGCTGCGCAGGCTTCTCATCTGGCCAGTCCACAGCAGCAGT 3521  
Qy 1192 IleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSerAsnThrGln 1211  
Db 3522 ATTTATCATGCGGGGCTGGCACCACCAACCATCTTCCATGACACCTGCTCTAATACACAG 3581  
Qy 1212 SerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHis 1231  
Db 3582 TCTCCAGAGCAGCTTTCCAGCAGCAGCAACAGCAGTCTTTCACCATCCACCTTCTCAT 3641  
Qy 1232 ValGlnProAlaTyrThrAsnProProHisMetAlaHisValProGlnAlaHisValGln 1251  
Db 3642 GTTCAGCGGGCATACACACCCACCCACATGCGCCACGTACCTCAGGCTCATGTACAG 3701  
Qy 1252 SerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMetThrThrGln 1271  
Db 3702 TCAGGAATGGTTCTTCTCATCCAATGCGCCATGATGCTTAATGACACACAG 3761  
Qy 1272 ProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProIleProValSer 1291  
Db 3762 CCACCC---GGTCCCAAGGCGCCCTCGCTCAAAAGTGCACTTACAGGCCATTCAGATTTTCG 3818  
Qy 1292 ThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaHisHisGlnGlnGln 1311  
Db 3819 ACAACAGCGCATTTTCCCTTATATGACGACCCCTTTCAGTACAAAGCCCAACCAACAGCAG 3878  
Qy 1312 Leu 1312  
Db 3879 TTG 3881

RESULT 10

AAV06553

ID AAV06553 standard; cDNA; 3798 BP.

XX

AC AAV06553;

DT 06-JUL-1998 (first entry)



QY 585 SerArgProSerArgProSerArgProProSerHisProSerAlaHisGlySerPro 604  
DB 1262 TCCAGGCCCTCCAGGCCATCCAGACCCCGCTCTCACCCCTCTGCTCATGGTTCTCCA 1321  
QY 605 AlaProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerPro 624  
DB 1322 GCTCTGCTCTACTATGCTTAAGCATGCTTTCAGAAAGNACCCCAAGGATGCTCCA 1381  
QY 625 LysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSer 644  
DB 1382 AAGGCACAGCGCCACCTCGGAATCACAGAGTCTCTGCTGGAGAGGCTCCATGCTAGT 1441  
QY 645 GlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArg 664  
DB 1442 GGCCTAGAAATTTGATCCCAATCCCAATCCCAATGAGCAGCTGCTCTCCAGTGCGCAAG 1501  
QY 665 ThrSerProSerGlyGlyThrTriSerSerValValSerGlyValProArgLeuSerPro 684  
DB 1502 ACCAGTCTCGAGGGGAACGTGCTCTCAGTGTCACTGGGTTCGAAGGTTATCTCCC 1561  
QY 685 LysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyPro 704  
DB 1562 AAAACTCACAGACCAGGTCTCCAGGCAGAGCAGCATTTGAAACTCTCCAGCGGGGCT 1621  
QY 705 ValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetProIlePro 724  
DB 1622 GTGCTGTGCTTCTCCCAAGCTGGCATCTCCCTGAGAAAGCGTTTCCATGCTGTTCCC 1681  
QY 725 AlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGlu 744  
DB 1682 GCCGATCTCCGACTCTCCAGCCCTGCATCCACAGAGCAGTGCACCCCTATTATGAG 1741  
QY 745 AlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsn 764  
DB 1742 GCAAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGGAGTAAAGAAAT 1801  
QY 765 IleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerPro 784  
DB 1802 GTTAAAGCAAGTGAACATCACTAGCTTTTCAAAGCTGACAAAGAGTATGTACCA 1861  
QY 785 ValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArg 804  
DB 1862 GTTGTCTTCGAACACAGAAACAGATTGATGACTTAAAGAGTTTAAAGATGATTTTAGG 1921  
QY 805 LeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGly 824  
DB 1922 TTACAGCCAAGCTCTACATCTGAATCTATGGATCAACTTAAAGCAAAATAGAGAAG 1981  
QY 825 GluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAspSerPheIle 844  
DB 1982 GAAAAGTCCAGAGATTTTGATTTAAAGATAAAAACGGAAGCAAGTGCTTAAGGATAGTTTCATT 2041  
QY 845 Glu-----AsnSerSerAsnCysThrSerGlySerSerLysProAsnSerProSer 862  
DB 2042 GACAGCAGCAGCAGCAGCAGCACTGTACAGTGGCAGCAGCAAGCAACAGCCCTAGC 2101  
QY 863 IleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGln 882  
DB 2102 ATCTCCCTTCCATGCTTAGTAAATGTCAGAGCACAAGAGGGGCTGAGGTCCATCCAA 2161  
QY 883 GlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluLysLys 902  
DB 2162 GGGGTGCAGATTCAGCCCAAGCTTGCAAAACAGAAAGAGGATGACAGAGAAGAGAGAA 2221  
QY 903 AspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnPro 922  
DB 2222 GACACAACAGAGCAGGTTAGGAATTCGACATTGAATCCCAATGCAAGAGTTCAACCT 2281  
QY 923 ArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGlnAlaGln 942  
DB 2282 CGTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACGTCACTCGGCCCTCAAGCAAA 2341  
QY 943 ProSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnProValCys 962

DB 2342 CCCAGCCCATCTATGGTGGGTTCATCAGCAGCAGCTCCAGTGTACACTCAGCCCTGTGTGC 2401  
QY 963 PheAlaProAsnMetMetTyrProValSerProGlyValGlnProLeuTyrPro 982  
DB 2402 TTCCGACCCCAATATGATGATCCCGTCCCAGTGAGCCCGGCGGTACAACCTTTATACCA 2461  
QY 983 IleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAla-----ValPro 1000  
DB 2462 ATACTATGAGCCCATGCTGTGAACCAAGCCAAGCATATAGAGCAGGTAAAGTACCA 2521  
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020  
DB 2522 AATATGCCCCAACAGCAGACCAAGCAACATCATCAAGCACCATGATGACCCAGCCTCC 2581  
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
DB 2582 GCGCAGGGCCACCANTCTGTAGCCACCCCGCCGCTTACTCCACTCAGTAGCTTGCCTTAC 2641  
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
DB 2642 AGCCTCAGCAGATTTCCTCAATCAGCCTTTTGGTCCAGCATGTGCCGCTTATCAGTCTCAG 2701  
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
DB 2702 CATCCTCATGTGTACAGTCTCTCATCAAGGTAAATGCCAGGATGATGGCACCAACCA 2761  
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
DB 2762 CATGCTCAGCTGGTTTAGTGTCTTCTTTCAGCTGCTCAGTTCGGGGCTCAGCAGCAGC 2821  
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
DB 2822 CACGCCATGTATGATGTCTCCCAAATTACCATAACAAGGAGACAAGCCCTTCTTCTTCTAC 2881  
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
DB 2882 TTTGCCATTTCCACCGGCTCCCTCGCTCAGCAGTATGCATCTCTTAATCGCCGCTGTCAT 2941  
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160  
DB 2942 CCACATCTCCCATCTCTCAGCCTTCGGCCACTCCACCGGACAGCAGCAAGCCAGCAT 3001  
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180  
DB 3002 GGTGGAAGTCACCTTCACCCAGTCTCTTTCAGCACCATCAGCAGCAGGCTGCCAGGCT 3061  
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
DB 3062 CTTTCATCTGGCCAGTCCACAGCAGCAGTGGGCCATTTATCATGCGGGCTGGCACCAACA 3121  
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
DB 3122 CCACCTTCCATGACACCTGCTCTTAATACAGTCTCCACAGACGAGTTTCCCAGCAGCA 3181  
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
DB 3182 CAACAGACAGTCTTCCACCATCCACCTTCTCATGTTCCAGCGGCATACACCCACCCACCC 3241  
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
DB 3242 CACATGGCCCCAGTACCTCAGGCTCATGTACAGTACAGAAATGGTTCCTTCTTCTCACTCA 3301  
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
DB 3302 GCCCATGGCCCAATCATGCTTAATGACACAGCAGCAGCC-----GGTCCCAAGCGCCCTC 3358  
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
DB 3359 GCTCAAAGTGCACGTACAGCCCATTCAGTTCGACAAACAGCGCATTTCCCTTATATGACG 3418  
QY 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312

Db	3419	CACCCCTTCAGTACAGCCACCACCAACAGCAGTTG	3454	
RESULT 11				
AAF21628				
ID	AAF21628	standard; DNA; 2006 BP.		
XX	AAF21628;			
AC	AAF21628;			
XX	27-MAR-2001	(first entry)		
DT				
XX				
DE		Human breast and ovarian cancer associated antigen gene SEQ ID 15.		
XX				
KW		Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;		
KW		neutropenic; neuroprotective; antiviral; antiallergic; hepatotropic;		
KW		antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;		
KW		antibacterial; antifungal; antiparasitic; cardiant; immune disorder;		
KW		Addison's disease; allergy; autoimmune haemolytic anaemia;		
KW		autoimmune thyroiditis; diabetes mellitus; Crohn's disease;		
KW		multiple sclerosis; rheumatoid arthritis; ulcerative colitis;		
KW		cardiovascular disorder; wound healing; neurological disease; ds.		
XX				
OS		Homo sapiens.		
XX				
PN		WO200055173-A1.		
XX				
PD		21-SEP-2000.		
XX				
PF		08-MAR-2000; 2000WO-US005881.		
XX				
PR		12-MAR-1999; 99US-0124270P.		
XX				
PA		(HUMA-) HUMAN GENOME SCI INC.		
XX				
PI		Rosen CA, Ruben SM;		
XX				
DR		WPI; 2000-611515/58.		
XX		P-PSDB; AAB58725.		
XX				
PT		New human breast and ovarian cancer associated gene sequences and the		
PT		polypeptides encoded by these genes, useful in the prevention, treatment		
PT		and diagnosis of cancer, immune disorders, cardiovascular disorders and		
PT		neurological diseases.		
XX				
PS		Claim 1; Page 490-491; 1299pp; English.		
XX				
CC		Sequences AAF21614 - AAF22031 represent DNA sequences encoding human		
CC		proteins AAB58711 - AAB59128. The DNA and protein sequences are		
CC		associated with breast and ovarian cancer. Included in the invention are		
CC		sequences AAF22032 - AAF22040 and AAB59129 which are used in the		
CC		isolation and characterisation of the DNA and protein sequences of the		
CC		invention. The breast and ovarian cancer associated DNA, protein, agonist		
CC		or antagonist sequences exhibit cytostatic; immunosuppressive; neutropic;		
CC		neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;		
CC		antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;		
CC		antifungal; antiparasitic and cardiant activity. The polynucleotide and		
CC		protein sequences are used in the diagnosis of cancer, particularly		
CC		breast and ovarian cancer. The nucleic acid sequences, proteins, agonists		
CC		and agonists may also be used in the diagnosis, prevention and treatment		
CC		of immune disorders e.g. Addison's disease, allergies, autoimmune		
CC		haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's		
CC		disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;		
CC		cardiovascular disorders such as myocardial ischaemias; wound healing;		
CC		neurological diseases such as cerebral anoxia and epilepsy; and		
CC		infectious diseases		
XX				
SQ		Sequence 2006 BP; 598 A; 577 C; 370 G; 456 T; 0 U; 5 Other;		
Alignment Scores:				
Pred. No.:	1,09e-65	Length:	2006	
Score:	2463.00	Matches:	460	
Percent Similarity:	98.72%	Conservative:	1	
Best Local Similarity:	98.50%	Mismatches:	2	
Query Match:	35.38%	Indels:	4	

DB:	3	Gaps:	1
US-10-802-228-2 (1-1312) x AAF21628 (1-2006)			
QY	850	AenCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSer	869
DB	6	AGTGTACCACTGGCAGCAGCAAGCGAATAGCCCGACATTTCCCTTCAATCTAGT	65
QY	870	AenThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerPro	889
DB	66	AACACGGAGCAACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACTTCCAGCCCA	125
QY	890	AlaCysLysGlnGluLysAspLysGluGluLysLysAspAlaAlaGlnValArg	909
DB	126	GCATGTAAACAGAGAAAGACGATGAAGAGAGAAAGAACCGCAGCTGAGCAAGTTAGG	185
QY	910	LysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLys	929
DB	186	AAATCAACATTGAATCCCAATGCAAGAGGTTCAACCCACCGTTCTTCTCTCAGCCAAAG	245
QY	930	ProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetValGly	949
DB	246	CCTTCTACTACCCCACTTCACTCGGCTCAAGCACAACCTAGGCCATCTATGGTGGT	305
QY	950	HisGlnGlnProThrProValThrGlnProValCysPheAlaProAsnMetMetTyr	969
DB	306	CATCAACAGCCCACTCCAGTTTATATCTAGCCCTGTTGTTTTCACCAAAATATGATGAT	365
QY	970	ProValProValSerProGlyValGlnProLeuTyrProIleProMetThrProMetPro	989
DB	366	CCAGTCCCAGTGGAGCCCGGCGTCAACCTTTATACCCAATACCTATGACGCCCATGCCA	425
QY	990	ValAsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGln	1007
DB	426	GTGAATCAAGCAAGACATATAGACAGCGTAAAGTACCAAAATATGCCCCAACAGCGGCA	485
QY	1008	AspGlnHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProIleAla	1027
DB	486	GACCAAGCATCACTCAGATGCCATGATGACCCAGCGTACAGCGGGCCCGGATTTGCA	545
QY	1028	AlaThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsn	1047
DB	546	GCMAACCCACAGCTTACTCCACGCAATATGTTGCTACAGTCCCTCAGCAGTCCCCAAT	605
QY	1048	GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerPro	1067
DB	606	CAGCCCTTGTTCAGCATGTGCCATATATAGTCTCAGCATCCTCATGTCTATAGTCTCT	665
QY	1068	ValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal	1087
DB	666	GTAATACAGGGTAATGCTAGATGATGGCACCACCAACACAGCCCGCTGGTTTAGTA	725
QY	1088	SerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysPro	1107
DB	726	TCTTCTTCAGCAACTCAGTACGGGGCTCATGACGACGACGATGCGATGATGATGATGCC	785
QY	1108	LysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySer	1127
DB	786	AAATTACCATACACCAAGGAGAGCAAGCCCTTCTTCTACTTTTGCCATTTCCACGGGCTCC	845
QY	1128	LeuAlaGlnGlnTyr-AlaHisProAsnAlaThrLeuHisProHisThrProHisProGln	1147
DB	846	CTTGCTCAGCAGTATGNGCRACCCCTAACCGTACCTCGCCCAACATCTCCACACCTCA	905
QY	1147	nProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGlyGlySerHisProAlaPr	1167
DB	906	GCCTTCAGTACCCCACTGGACAGCAGCAAGCAACCAACCTGGTGGAGTCTCTCTGACC	965
QY	1167	oSerProValGlnHisGlnHisGlnAlaAlaGlnAla-LeuHisLeuAlaSerProGln	1187
DB	966	CAGTCTCTGTTACGACACATCAGCAGCCGCGCCAGGCTTCTCCATCTGCGCCAGTCCAC	1025
QY	1187	lnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProA	1207









```
QY 948 aGlyHisGlnProThrProValThr-----GlnProValCysPheAlaProA 966
Db 2397 -----CGGTCTGACAGCAGCCGAGAGTGGCTATACAGCCGCC 2436
QY 966 smMetMetTyrProValPro-----ValSerProGlyValGln-----ProLeuT 981
Db 2437 AGTACATCTCTTACATACCTCAGATCCATGGGACGAGCTGTGAGGACACCTCAGATG 2496
QY 981 YrProleProMetThr-----ProMetProValAsnGlnAlaIleYrThrYrAlaValP 1000
Db 2497 ATCCATATCTGTATCCATTCAGTTCCTGGGAGCAGGCAAGTACCGGGGAGCAAAAG 2556
QY 1000 roAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAla-----MetMetH 1017
Db 2557 GCTCCCTCTCTCGCAGCGCTCGGACCAACACCGACCGCCTCAGCCCGCCGCGATGATGC 2616
QY 1017 isProAlaSerAlaAlaGlyProProleAlaAlaThrProProAlaTyrSerThrGlnT 1037
Db 2617 AGCCCGCGCGGTCTGTGCGCCCTCTGTGTGCTGCCAGGCC-----TATTCTTC--T 2670
QY 1037 YrValAlaTyrSerProGlnGlnPheProAsnGlnPro-----LeuValGlnHisValProH 1056
Db 2671 ACATCCCTTACAACTCAGCAGTCTCCAGGCCAGCCATGATGACGCCCATGGGCC 2730
QY 1056 isTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetM 1076
Db 2731 ACTACCCCTCACAG-----CCGCTGTTTGGCCCATGTCTCAGACCAACCCAGCGATGC 2784
QY 1076 etAlaProProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyA 1096
Db 2785 TGACGTGGCGGCACCATCCCGCAG-----GCCATCGTGTATCTCTACCCCTCAATACCTT 2841
QY 1096 laHisGluGln-----ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysG 1114
Db 2842 CTGCAGAGAGCGCTACCCCAAGCCCTTATGCC----- 2876
QY 1114 luThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaH 1134
Db 2877 -----ACTGTTCCACGATCTCATCCAC 2898
QY 1134 isProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrG 1154
Db 2899 ACCATGCCACAGAGCTCCATGCCAC-----CAGCGCGAGCGGCTTACCACCCCTACTG 2952
QY 1154 yGlnGln-----GlnSerGlnHisGlyGlySerHisProAlaProSerProValGlnHis- 1172
Db 2953 GAAGCAGCGCGAGTCCAGCATGCG-----GCCCCAGTCTGTGTCCAGT-G 2999
QY 1172 ----- 1172
Db 3000 CCTGCCATGGGGGTGCTGAGTGTCTCTGTGTCAGGAATGGTGGCCAGAGAAGGATA 3059
QY 1172 ----- 1172
Db 3060 GAGCTAGGGGTCAATTTCTGAGTGGCGAGGACTGGGGCCAGCAGTGTGTGGCTGTGTG 3119
QY 1173 -----HisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerP 1186
Db 3120 GCACTCAACCTTCCCTCCCGCAGCAGCATCAGGGGGGAGGCCCCACACTTGGGCAGTG 3179
QY 1186 roGlnGlnGlnSerAlaIleTyrHisAlaGly-----LeuAlaProThrProProSerMetT 1205
Db 3180 GACAGCCACAGCAAGAACTGTATCCACCCAGGGGCGCTTGACAGGCAGCCGCGCTCTCTGC 3239
QY 1205 hrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValP 1225
Db 3240 CACCGGAGCTTCTGCGCCAGTCCCTCAGAGAGCTTCCCTC-----CAGCCAGCGCGTGTGT 3296
QY 1225 heThrIleHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisV 1245
Db 3297 ATGCCATCCACACAGCAGCTGCCCGCAGCGCTTCAACCAAC-----ATGCCCATG 3347
QY 1245 alProGlnAlaHisValGlnSerGlyMetVal-----ProSerHisProThrAlaH 1262
```

```
Db 3348 TTACCAGCCCATGTCCAACTGGAATCAGCAGAGCCCGCCCTCACCTCGGGCTC 3407
QY 1262 isAlaProMetMetLeuMetThrThrGlnProPro-----GlyGlyProGlnA 1278
Db 3408 CCCACCCGCCCGCAGGTGATGTGTGTCACCCACCCAGAGTCATGGGGGGCCCCCAAG 3467
QY 1278 laAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProT 1298
Db 3468 GCGCGGTGCCCGCAGAGTGGGTGCTGCACTCTCAGCTTCCACACCTCACCTACCCAT 3527
QY 1298 YrMetThrHisPro 1302
Db 3528 ACATCGGACACCCC 3541

RESULT 14
ABZ72015
ID ABZ72015 standard; cDNA; 4674 BP.
XX
AC ABZ72015;
XX
DT 01-APR-2003 (first entry)
XX
DE Human ataxin-2-like protein A2LP (A2LG) cDNA GenBank AF034373.
XX
KW Human; cancer; stomach cancer; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283899-A1.
XX
PD 24-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-JP003038.
PR 10-APR-2001; 2001JP-00112039.
PR 21-SEP-2001; 2001JP-00290193.
XX
PA (TAKA-) TAKARA BIO INC.
XX
PI Yoshikawa Y, Okamoto S, Oura T, Mineno J, Asada K, Kato I;
PI Inoue H, Mori M;
XX
DR WPI; 2003-093022/08.
XX
PT Measuring changes in expression of 264 cancer associated genes for
PT detection of stomach cancer and screening of potential anticancer agents.
XX
PS Claim 2; Page; 266pp; Japanese.
XX
CC The invention relates to a method for the detection of cancer in which a
CC change in the expression of 1 or more of 264 specified cancer associated
CC genes, ABZ71694-ABZ71957, or of sequences at least 80% homologous to them
CC in the specimen tissue as compared to normal tissue is observed. The
CC genes are used in detection, diagnosis and treatment of cancer,
CC especially of stomach cancer. The present sequence is that of a cancer
CC associated polynucleotide of the invention. Note: The present sequence
CC was not given in the printed specification but was isolated using the
CC GenBank accession number given in the DE line
XX
SQ Sequence 4674 BP; 971 A; 1566 C; 1244 G; 893 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,58e-46 Length: 4674
Score: 1816.50 Matches: 504
Percent Similarity: 48.22% Conservative: 174
Best Local Similarity: 35.85% Mismatches: 381
Query Match: 26.10% Indels: 350
DB: 8 Gaps: 52

US-10-802-228-2 (1-1312) x ABZ72015 (1-4674)
QY 19 ArgPheAlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSer 38
```

Db	6	CGCTCCTCAC	TTCCTAGATGGGATCGCGCGGGAAGAGCGCTCCTCATTTTCTTAGATG	65
Qy	39	GlyArgGlyGly	GlyGlyAlaProGlyProTyrProSerAlaAlaProProProProProProProPro	58
Db	66	GGATGGCGCGCGG	CAGAGATGCTCTCACTTCTCAGCGCGAGACCCCTCTCCCTT	125
Qy	59	GlyProGlyProProPro	SerArgGlnSer-----SerProProSerAlaSerAspCys	76
Db	126	CGCGCTCGCGCGCT	TCTCTCGCGCGGTTCCTCTCCACCC-----	170
Qy	77	PheGlySerAsnGly	AsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGly	96
Db	171	-----	CGCACCCGCG-----	182
Qy	97	LeuGlyGlyProProArg	ProPheValValValLeuLeuLeuAlaSerProGlyAla	116
Db	183	-----	GGTCCCCCGCGCG-----	218
Qy	117	ProProAlaAlaProThr	ArgAlaSerProLeuGlyAlaArgAlaSerProProArgSer	136
Db	219	CCGATCCCCCTCGCT	TCCGCGC-TCTCAGCGGGCCCCAGCCCCCTCTCTCTCC	277
Qy	137	GlyValSerLeuAlaArg	ProAlaProGlyCysProArgProAlaCysGluProValTyr	156
Db	278	CTCCCTTCTCTCTA	-ATTCCCTTCCGACGCT-----	309
Qy	157	GlyProLeuThrMetSer	LeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	176
Db	310	-----	GGCATCATGTTGAAGCTCAGCGCTACCAACAGCCCTCCAGCGCCCGAGCAG	360
Qy	177	GlnGlnGlnGlnGlnGln	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	196
Db	361	CCGCCCCCACGCAAT	CGCTGCGCGTCGGCCCCCGGGGGCACGCCCTCCCAACG	420
Qy	196	GlyAspProGlyGly	-----SerGlyLeuLeu	204
Db	421	CGCGCTCCCGGGCG	GCTGGCCACTCTGCGGCTCTCCCGGGCTCTCCAGCGCGCGCT	480
Qy	204	uAlaSerProAlaAla	-----ProSerProSerSerSerSerValSe	219
Db	481	CCCCCTGCTGGGCG	CTGTGGCGTGGCGGAGCGGGCTCCCGCGGGAGCGGAGCGCA	540
Qy	219	rSerSerSerAlaThrAla	ProSerSerValValAlaAlaAlaThrSerGlyGlyGlyArgPr	239
Db	541	TCTTGGCGCGCAGCG	CGCGCGCGACACACAGGAGAGCGCGGGGCGAGCGCGCA	600
Qy	239	o-GlyLeuGlyArgGly	ArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheA	259
Db	601	TCGGCAGCGCCAGGG	CAGAGCACAGGAAAGGACCCCCACAGTCACCTGTG	657
Qy	259	spGlyIleTyrAlaAsn	MetArgMetValHisIleLeuThrSerValValGlySerLysC	279
Db	658	AAGCGCTCTACAACA	ATTCAGAAATGCTGCAATTCCTTACAGCTGTGTGGGCTCCACTT	717
Qy	279	ysGluValGlnValLys	AsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProL	299
Db	718	GTGATGTAAAGGTGA	AAATGGTACCACCTTATGAGGGTATCTTCAAGACGCTAAAGCTCA	777
Qy	299	ysCysAspLeuValLeu	AspAlaAlaHisGluLysSerThrIleSerSerGlyProL	319
Db	778	AGTTTGAAGTACGCG	TGTGTGCACCGGAAAGCATCTGAGCGAGCGGTCGCCCTC	837
Qy	319	ysArgGluGluIleMet	clSerIleLeuPheLysCysSerAspPheValValGlnP	339
Db	838	GTCCGGGAGGACAT	TGTGGACACCAATGGTGTAAAGCAAGTTCACCGATTCAGCCATTGCCA	897
Qy	339	heLysAspMetAspSer	SerTyrAlaLysArgAspAlaPheThrAspSerAlaIle---	357
Db	898	TCCGAAATGTTGACT	TTCACCTATGCTACTAAGACAAGTTCACCGATTCAGCCATTGCCA	957
Qy	358	--SerAlaLysValAsn	GlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyG	377

QY 727 -----SerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluA 745  
DB 1777 TGGGCTCTCATTTCTCCAGCTTCTCCAAAGATCTCTCCGCGCCCAAC----- 1823  
QY 745 laLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnI 765  
DB 1824 -----GATGTAAAGAACTCTTACCNAAGAACTCTGGGAACTC 1863  
QY 765 leLysProAsnGlnThrSerProSerPheSerLysAlaGluAsnLysGlyLysSerProV 785  
DB 1864 TGGAGCCCGCAGGAGCTG-----GCTCGGATAGCTGGGAAGTCCCTG 1905  
QY 785 alValSerGluHisArgLys-----GlnIleAspAspLeuLysLysPheLysAsnAspP 803  
DB 1906 GTCTTTCAGAAATGAACAGAACGATTCCTCACTGGAAGAACTGAGAAAGTTTGGGGCCCACT 1965  
QY 803 heArgLeuGlnProSerSerThrSerGlu--SerMetAspGlnLeuLeuAsnLysAsnA 822  
DB 1966 TTAAGCTTCAGCCCACTAGTCTCCCTGGAACAGCCTGGATCCTTCTTCTCCCGG----- 2021  
QY 822 rgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAspS 842  
DB 2022 -----ATCTTAAGGAG-----GAGCCCAAGGAAGGAGA 2052  
QY 842 erPheIleGluAsn-----SerSerSerAsnCysT 852  
DB 2053 AAGAGCTGATGCTGTCTGTTGACTTCAGAGCCCATGGGGTCTCCCGTCTCTCCAAAGACAG 2112  
QY 852 hrSerGlySerLysProAsnSerProSerLysSerProSerLysLeuSerAsnThrG 872  
DB 2113 AGTCCGTATCGGATAGGAGGACAAACACCTCTGGCACATCA-----GGAGGCACATG 2166  
QY 872 luHisLysArgGlyProGluValThrSerGlnGlyVal-----GlnThrSerSerProA 890  
DB 2167 AG-----GGCCAGAGCAGCCCGCCACACCTTGTCCAAAGCCAACTGGCAGCCGCC 2217  
QY 890 laCysLysGlnGluLysAspLysGluGluLysLysAspAlaAlaGluGlnValArgL 910  
DB 2218 CGGTGGGCTCATCAAGGAGGAGAGACAAAGATGAGGGCCCTGTTGTGAACAAGTAAAGA 2277  
QY 910 ysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSer-----PheSerGlnP 928  
DB 2278 AATCAACGTTGAACCCCTAATGCTTAAGGAGTTCAATCTTACAAAGCCTCTGTGTCTGTGA 2337  
QY 928 roLysProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetV 948  
DB 2338 ATAAATCCACCGAGTACCCCACTTCTCCGGGCCCCGGACTCATTTCAACTCTCCCTCATC- 2396  
QY 948 alGlyHisGlnGlnProThrProValThr-----GlnProValCysPheAlaProA 966  
DB 2397 -----CGGTGCTGACAGCAGCCAGAGTGGGCTATACAGCCGCC 2436  
QY 966 snMetMetTyrProValPro-----valSerProGlyValGln-----ProLeuT 981  
DB 2437 AGTACATCTCTACATACCTCAGATCCATCGGAGCAGCTGTGAGGACCTCAGATGT 2496  
QY 981 yrProIleProMetThr---ProMetProValAsnGlnAlaLysThrTyrArgAlaValP 1000  
DB 2497 ATCATATCTGTATCAATTCATGCTGGGAGCAGGCAAGTACCGGGGAGCAAAAG 2556  
QY 1000 roAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAla-----MetMetH 1017  
DB 2557 GCTCCCTCTCTCCGAGCGCTCGGACCAACACAGCCAGCCTCAGCCCGCGCGATGATGC 2616  
QY 1017 isProAlaSerAlaAlaGlyProProIleAlaThrProProAlaLysThrGlnT 1037  
DB 2617 AGCGCCGCGCGGTGTGTGGCCCGCTCTGTGTGGTGTGCGACGCC-----TATTCTTCC--T 2670  
QY 1037 yrValAlaLysSerProGlnGlnPheProAsnGlnPro---LeuValGlnHisValProH 1056  
DB 2671 ACATCCCTTACAAACCTCTCAGAGTTTCCAGCCAGCCAGCCATGATGATGAGCCCATGGCCC 2730

QY 1056 isTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetM 1076  
DB 2731 ACTACCTCTCAG-----CCGGTGTGTCCTCCATGCTTCAGAGCAACCCACCGCATGC 2784  
QY 1076 etAlaProProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyA 1096  
DB 2785 TGAGCTCGGGCAGCATCCCCAG---GCCATCGTGTCTCTACCTCTCAATACCTCTT 2841  
QY 1096 laHisGluGln-----ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysG 1114  
DB 2842 CTGAGCAGCAGCCTACCCCAAGCCCTTATGTC----- 2876  
QY 1114 luThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaH 1134  
DB 2877 -----ACTGTTACCCAGTCTTACCAC 2898  
QY 1134 isProAsnAlaThrLeuHisProHisThrProHisProGlnProSerSerAlaThrProThrG 1154  
DB 2899 ACCATGCCACACAGCTCCATGCCAC-----CAGCCGACGCGGCTACCCGCTACTGT 2952  
QY 1154 lyGlnGln---GlnSerGlnHisGlyGlySerHisProAlaProSerProValGlnHis- 1172  
DB 2953 GAAGCCAGCCGAGTCCAGCATGCG-----GCCCCAGTCTGTGTCAGGT-G 2999  
QY 1172 ----- 1172  
DB 3000 CTGTCATGGGGGTGTGAGTGTCTGTGTCAGGAATGGTGCCAGAAAGGGATA 3059  
QY 1172 ----- 1172  
DB 3060 GAGCTAGGGGTCAATTTCTGAGTGGCGAGACTGGGGGCCAGCGAGTTGCTGCTGTGTG 3119  
QY 1173 -----HisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerP 1186  
DB 3120 GCACCTAACTTCCCTCTCCCGCAGCAGCATCAGCGGGGCGGCGCCACACTTGGGCAGTG 3179  
QY 1186 roGlnGlnSerAlaIleTyrHisAlaGly---LeuAlaProThrProProSerMetT 1205  
DB 3180 GACCCACAGCAGAAATCTGTATCCACCCAGGGGCCCTGACAGGACGCGCCCTCTCTGTC 3239  
QY 1205 hrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValP 1225  
DB 3240 CACCGGGACCTTCTCCCGCAGTCCCTCAGAGCAGCTTCCCC---CAGCCAGCGCTGTCT 3296  
QY 1225 heThrIleHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisV 1245  
DB 3297 ATGCCATCCACCCAGCAGCTGCCCGCCAGCGCTTCCCAAC-----ATGGCCCATG 3347  
QY 1245 alProGlnAlaHisValGlnSerGlyMetVal-----ProSerHisProThrAlaH 1262  
DB 3348 TTACCCAGGCCATGTCCAACTGGAATCAGAGCAGCCCGCCCTCACCCTGGGGCTC 3407  
QY 1262 isAlaProMetMetLeuMetThrThrGlnProPro-----GlyGlyProGlnA 1278  
DB 3408 CCCACCGCCCGCAGGTGATGTGTGTCACCCACCCAGAGTCTATGGGGGCCCCCCCCAAG 3467  
QY 1278 laAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProT 1298  
DB 3468 GCGGCTGCCCGCAGAGTGGGGTGCCTGCTCTCAGCTTCCACACCTTCACTCCATCCCAT 3527  
QY 1298 yrMetThrHisPro 1302  
DB 3528 ACATCGACACCC 3541  
RESULT 15  
ADH29078  
ID ADH29078 standard; DNA; 4674 BP.  
XX  
AC ADH29078;  
XX  
DT 11-MAR-2004 (first entry)  
XX Human chronic myelogenous leukaemia (CML) gene marker #346.  
DE



476 snSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnIleValProProG 496  
1312 AGGGCTCAGGGCGGAGAGCCCGAGCTTGGCATCCAGGGAGGAGTATATCCCTCTGC 1371  
496 lylGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetG 516  
1372 CTCAACGAGTCCGGGAGGT---CCCCGGGAGGAGTTCGATGCACGACTCTCGGGGCG 1428  
516 lylGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAspPheAsnP 536  
1429 GTGGGCTGGCCCTAGCTCTTTGCCACCTCGTGGCCCTCACCATCTGGACAACAGCAGCC 1488  
536 roAsnSerGlySerAspGlnArgValIleAsnGlyValProTrpProSerProCysP 556  
1489 CTGGCCAGGTTCTGAGGCCGCGTGTATCAATGGA----- 1523  
556 roSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProA 576  
1523 ----- 1523  
576 rgAlaAlaThrProThrArgProProSerArgProProSerArgProProS 596  
1523 ----- 1523  
596 erHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMetSerS 616  
1523 ----- 1523  
616 erGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValS 636  
1524 -----GGCCCTTCCCGCATGTCCTCCAAAGGCACAGCGGCTCTGAGA----- 1565  
636 erAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSerG 656  
1566 -----GGTGCCAGACTCTCTCTCG-----CCAGTA 1593  
656 luAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerValV 676  
1594 ATAGGCTTCTGGAGAACTCTGTGCCACCTCTCTCGAGTGGG----- 1637  
676 alSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnS 696  
1638 -----CCGGATGTATCCCGCGTCTCCCAAGACTGTGTCGCCCT----- 1676  
696 erIleGlyAsnThrProSerGlyProValLeuAlaSer-----ProGlnAlaG 712  
1677 -----GCCCAATCTCAGCTTCTCTCCAGCCTCTCTCCAGCCTCCCATCGGCT 1716  
712 lylIleIleProThrGluAlaValAlaMetProIleProAlaAla----- 726  
1717 CGGCAGTGCCAACTCTTCAGCCTCCATCCCTGTGACCTCATCAGTCTCAGATCTCGAG 1776  
727 -----SerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluA 745  
1777 TGGGCTCCATTTCTCCAGCTTCTCCAAAGACTCTCCCTGGGCCCCACACA----- 1823  
745 lalysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnI 765  
1824 -----GATGTAAAGAACTCTCTTACCAAGAACTCTGGGAGAACTC 1863  
765 leLysProAsnGluThrSerProSerPheSerIysAlaGluAsnLysGlyIleSerProV 785  
1864 TGGAGCCCGAGGACTG-----GCTCGGATAGTGGGAAAGTCCCTG 1905  
785 alValSerGluHisArgLys-----GlnIleAspAspLeuLysLysPheLysAsnAspp 803  
1906 GTCTTCAGATGAACAGAAACGATTCCAACTGGAGAACTGAGAAAGTTTGGGGCCCACT 1965  
803 heArgLeuGlnProSerSerThrSerGlu-----SerMetAspGlnLeuLeuAsnLysAsnA 822  
1966 TTAAGCTTCAGCCAGTAGCTCCCTCAGAACACGCTGGATCTTTCTCCCGCG----- 2021  
822 rgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAsps 842

2022 -----AICTTAAAGAG-----GAGCCCAAGAAAGGAGA 2052  
842 erPheIleGluAsn-----SerSerSerAsnCysT 852  
2053 AAGAGTTGAGTGGTCTGTTGACTTCAGAGCCCATGGGTCTCCCGTCTCTCTCCAGACAG 2112  
852 hrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThrG 872  
2113 AGTCCGTATCGATAGGAGGACAAACACCCTCGCACCATCA-----GGAGGCACTG 2166  
872 lylHisLysArgGlyProGluValThrSerGlnGlyVal-----GlnThrSerSerProA 890  
2167 AG-----GGCCAGAGCAGCCCCCACCCTTGTCCAAGCCAACTGGCAGGCCCC 2217  
890 laCysLysGlnGluLysAspLysGluLysLysAspAlaAlaGluGlnValArgL 910  
2218 CGGTGGGCTCATCAAGGAGAGAAACAAGATGAGGGCCCTGTTGCTGAACAAGTAAAGA 2277  
910 ysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSer-----PheSerGlnP 928  
2278 AATCAACGTTGAACCTTAATGCTAAGGAGTTCAATCTCAAAAGCTCTGCTGTCTGTGA 2337  
928 roLysProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetV 948  
2338 ATAAATCCACAGTACCCCAACTTCTCCGGGCGCCCGGACTCATTCACACTCCCTCCATC- 2396  
948 alGlyHisGlnGlnProThrProValTyrThr-----GlnProValCysPheAlaProA 966  
2397 -----CCGGTCTCACAGCAGCCAGAGTGGGTATACAGCCCCC 2436  
966 snMetMetTyrProValPro-----ValSerProGlyValGln-----ProLeuT 981  
2437 AGTACATCTCTACATACATCCAGTCCAGTGGAGCAGCTGTGCAGGCACCTCAGATGT 2496  
981 yrProIleProMetThr-----ProMetProValAsnGlnAlaLysThrTyrArgAlaValP 1000  
2497 ATCCATATCTGTATCCAAATTCAGTGCCTGGCAGCAGGCAAGTACCGGGAGCAAAAG 2556  
1000 roAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAla-----MetMetH 1017  
2557 GCTCCCTTCTCCGAGCGCTCGGACCAACACCAGCAGCCTCAGCCCCCGCGATGATGC 2616  
1017 isProAlaSerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnT 1037  
2617 AGCGCCCGCGGCTGTGCGCCGCTCTGTGTGCTGCCAGCC-----TATTCTTCC---T 2670  
1037 yrValAlaTyrSerProGlnGlnPheProAsnGlnPro-----LeuValGlnHisValProH 1056  
2671 ACATCCCTTACAACTCAGCAGTTCAGGCGCAGCCAGCCATGATGACGCCCATGGCCCC 2730  
1056 isTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetM 1076  
2731 ACTACCCCTCACAG-----CCGGTGTGCCCCCATGCTTTCAGACCAACCCAGCATGC 2784  
1076 etAlaProProThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyA 1096  
2785 TGAGCTGGCGCAGCCATCCCCAG---GCCATCGTGTCTCTCTACCCCTCAATACCTT 2841  
1096 laHisGluGln-----ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysG 1114  
2842 CTGAGAGCAGCTACCCCAAGCCCTTTATGCC----- 2876  
1114 luThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaH 1134  
2877 -----ACTGTTCCAGCTCTTACCAC 2898  
1134 isProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrG 1154  
2899 ACATGCCACACAGCTCCATGCCAC-----CAGCGCAGCGGCTTACCAGCCTACTG 2952  
1154 lylGlnGln-----GlnSerGlnHisGlyGlySerHisProAlaProSerProValGlnHis- 1172



```

Db      2953 GAAGCAGCGCGAGTCCCGAGATGG-----GCCCCAGTCTGTCCAGGT-G 2999
Qy      1172 ----- 1172
Db      3000 CCTGCATGGGGGTGCTGAGTGGTCTTGGTGCAGGAATGGGTGGCCAGAGAAGGGATA 3059
Qy      1172 ----- 1172
Db      3060 GAGTAGGGGTCAATTTCTGAGTGGCGAGACTGGGGGCCAGCGAGTTGCTGGCCTGTGTG 3119
Qy      1173 -----HisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerP 1186
Db      3120 GCACCTCAACCTTCCCTCCCGAGCGATCAGCGGGGGGCGGCCCCACACTTGGGCAGTG 3179
Qy      1186 roGlnGlnSerAlaIleTyrHisAlaGly---LeuAlaProThrProProSerMetT 1205
Db      3180 GACAGCCACAGCAGAACTGTACACCCAGCGGGCCCTGACAGGCACGCCGCCCTCTCTGC 3239
Qy      1205 hrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValP 1225
Db      3240 CACCGGGACCTTCTGCCCGAGTCCCTTCAGAGCAGCTTCCCC---CAGCCAGCGCGTGTGT 3296
Qy      1225 heThrIleHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisV 1245
Db      3297 ATGCCATCCACCACAGCAGCTGCCCGCGGCTTACCACAC-----ATGGCCCATG 3347
Qy      1245 alProGlnAlaHisValGlnSerGlyMetVal-----ProSerHisProThrAlaH 1262
Db      3348 TTACCCAGGCCCATGTCCAACTGGAATCACAGCAGCCCCCGCCCTCACCTTGGGGCTC 3407
Qy      1262 isAlaProMetMetLeuMetThrThrGlnProPro-----GlyGlyProGlnA 1278
Db      3408 CCCACCGCGCCCGAGTGATGTGTGTCACACCCACCCAGAGTCATGGGGGGCCCCCAAG 3467
Qy      1278 laAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProT 1298
Db      3468 GCGCGGTGCCCCAGAGTGGGGTGCCTGCACTCTCAGCTTCCACACCCCTACCCATACC 3527
Qy      1298 yrMetThrHisPro 1302
Db      3528 ACATCGGACACCCC 3541

```

Search completed: December 23, 2005, 06:53:42  
Job time : 1770 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 23, 2005, 05:15:22 ; Search time 9990 Seconds  
(without alignments)  
6144.610 Million cell updates/sec

Title: US-10-802-228-2  
Perfect score: 6961  
Sequence: 1 MRSAAAAPSPAVATESRRF.....TAHFPMTWPSVQAHHQOOL 1312

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US10802228/runat\_21122005\_110225\_9713/app\_query.fasta\_1.1479  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10802228 @CGN 1 1 8649 @runat\_21122005\_110225\_9713 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*  
10: gb\_est10: \*  
11: gb\_est11: \*

CR619107 1528 bp mRNA linear HTC 21-JUL-2004  
full-length cDNA clone CS0D1087Y117 of Placenta Cot 25-normalized  
of Homo sapiens (human).  
CR619107 GI:50499914  
HTC; CDSLT\_CDNA.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 1528)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
2 (bases 1 to 1528)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2106.5	30.3	1528	4 CR619107	CR619107 full-length
2	1644.5	23.6	1060	1 AL554896	AL554896 AL554896
3	1589	22.8	977	7 CN646951	CN646951 ILLUMIGEN
4	1556	22.4	912	5 BQ883680	BQ883680 AGENCOURT
5	1530	22.0	938	5 BX375946	BX375946 BX375946
6	1525	21.9	876	3 BP159935	BP159935 BP159935
7	1508.5	21.7	1000	5 BX402110	BX402110 BX402110

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

FEATURES		Location/Qualifiers	
source		1..1528	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CSODI087Y117"	
		/tissue_type="Placenta Cot 25-normalized"	
		/plasmid="pCMVSPORT_6"	
ORIGIN			
Alignment Scores:			
Pred. No.:		3.85e-100	Length: 1528
Score:		2106.50	Matches: 395
Percent Similarity:		92.18%	Conservative: 6
Best Local Similarity:		90.80%	Mismatches: 19
Query Match:		30.26%	Indels: 15
DB:		4	Gaps: 3
US-10-802-228-2 (1-1312) x CR619107 (1-1528)			
QY	861	ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr	880
DB	1	CCACGCAITTCCTCCCTCAATACCTTAGTAACAGGAGCACAGAGGGGACCTGAGGTCACT	60
QY	881	SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu	900
DB	61	TCCCAAGGGGTTTCAGATTCAGCCCGCAGCATGTAAACAAGAGAAAGACGATAAGGAAGAG	120
QY	901	LysLysAspAlaAlaGluGlnValArgLysSerThrIleuAsnProAsnAlaLysGluPhe	920
DB	121	AAGAAAGACGCGCTGAGCAAGTTAGGAATCAACATTTGAATCCCAATGCCAAGAGGATTC	180
QY	921	AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln	940
DB	181	AAACCAAGTTCCTCTCTCAGCCAAAGGCTTCTACTACCCCAACTTCACCTCGGCTCAA	240
QY	941	AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro	960
DB	241	GCACAACCTAGCCCATCTATGGTGGTCAATCAACAGCCAACTCCAGTTTATATCTCAGCCT	300
QY	961	ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980
DB	301	GTTTGTGTTGCAACCAAAATATGATGATCCAGTCCAGTCCAGGCCAGCCCGGCTCAACCTTTA	360
QY	981	TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAla-----	998
DB	361	TACCCAAATACCTATGACGCCCATGCGAGTGAATCAAGCCNAGACATATAGACGAGTAAA	420
QY	999	ValProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisPro	1018
DB	421	GTACCAAAATATGCCCAACAGCGGCAAGACCAAGCATCATCAGATGCCATGATGCACCCA	480
QY	1019	AlaSerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrVal	1038
DB	481	CGGTTCAGCAGCGGGCCCAACCGATTCAGCCACCCACCCAGCTTACTCCACGCAATATGTT	540
QY	1039	AlaTyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGln	1058
DB	541	GCCTACAGTCTTCAGAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAG	600
QY	1059	SerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaPro	1078
DB	601	TCTCAGCATCTCATGTCTATAGTCTGTATACAGGGTAAATGCTAGAAATGATGCACCA	660
QY	1079	ProThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGlu	1098
DB	661	CCAAACACAGCCCGCCCTGGTTTATGATCTCTTTCAGCAACTCAGTACGGGCTCATGAG	720
QY	1099	GlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSer	1118
DB	721	CAGAGCGCATGCATGTATGCATGTCTCCAAATTTACCATACACAGGAGACAAAGCCCTTCT	780
QY	1119	PheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThr	1138
DB	781	TTCTACTTTGCCATTTCCACGGGCTCCTTGTGCTCAGCAGTATGCGCACCTTAAGCTACC	840
QY	1139	LeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSer	1158
DB	841	CTGCACCCACATACTCCACACCTCAGCCTTACGCTTACCTCCCTGACACAGCAAGCAAGC	900
QY	1159	GlnHisGlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAla	1178
DB	901	CAACATGGTGAAGTCATCTGCACCCAGTCTCTGTTTCAGCACCATCAGCACCAGGCGGCC	960
QY	1179	GlnAlaLeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAla	1198
DB	961	CAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCAGCAATTTACCAACGCGGGCTTGGC	1020
QY	1199	ProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhePro	1218
DB	1021	CCAACTCCACCTCCATGACACCTGCTTCCAAACAGCAGTCCGACAGATAGTTTCCCA	1080
QY	1219	AlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsn	1238
DB	1081	GCAGCAACAACAGACTGCTTTACGATCCATCTCTTCACGTTTCAGCGCGGTATACCAAC	1140
QY	1239	ProProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHis	1258
DB	1141	CCACCCCACTGGCCCGCAGTCAGTACCTCAG-----TACAAG	1173
QY	1259	ProThrAlaHisAlaProMetMetLeuMetThrThrGlnProGlyGlyProGlnAla	1278
DB	1174	CCCACCAACCAACAGTTGTAAGGTCGCTTGAAGGAGGAGCAACCAAGGC-----	1221
QY	1279	AlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThr 1293	
DB	1222	CAATTCCTCCTCCCTCTACTGCTTCTACCAACTGGAAGCACA 1266	
RESULT 2			
AL554896			
LOCUS			
DEFINITION			
AL554896 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
clone CSODI087Y117 5-PRIME, mRNA sequence.			
ACCESSION			
AL554896.3 GI:45859645			
VERSION			
EST.			
KEYWORDS			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 1060)			
AUTHORS			
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.			
TITLE			
Full-length cDNA libraries and normalization			
JOURNAL			
Unpublished (2001)			
COMMENT			
On Feb 15, 2001 this sequence version replaced gi:31276706.			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime			
end enriched, double-strand cDNA was digested with NotI and cloned			
into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library			
was normalized. Library was constructed by Life Technologies, a			
division of Invitrogen. This sequence belongs to sequence cluster			
6451.r			
For more information about this cluster, see			
http://www.genoscope.cns.fr/cdna?s=CSODI087AE09QPI&c=6451.r.			
FEATURES			
source			
Location/Qualifiers			
1..1060			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CSODI087Y117"			
/tissue_type="PLACENTA COT 25-NORMALIZED"			
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"			





```

QY 787 SerGluHisArgLysGlnIleAspLeuLysLysPheLysAsnAspPheArgLeuGln 806
Db 241 TCTGAACATAGAAAACAGATTCATGATTTAAAGAAATTTAAGAAATGATTTAGGTTACAG 300
QY 807 ProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyGluLys 826
Db 301 CCAGGTTCTACTTCTGAATCTATGGATCAACTACTTAACAAAATAGAGGGAGNAAA 360
QY 827 SerArgAspLeuLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsn 846
Db 361 TCAAGAGATTTGATCAACAGACAAAATGAACCAAGTGCTAAGGATCTTTTCATTGAAAT 420
QY 847 SerSerSerAsnCysThrSerGlySerSerLysProAsnSerProSerIleSerProSer 866
Db 421 AGCAGCAGCAACTGTACCAAGTGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 480
QY 867 IleLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThr 886
Db 481 ATACTTAGTAAACAGGAGCAAGAGGGACCTGAGTCACTTCCCAAGGGTTCAAGT 540
QY 887 SerSerProAlaCysLysGlnGluLysAspLysGluLysLysAspAlaLys 906
Db 541 TCCAGCCCAGCATGTAACAAAGAGAAAGACGATAAGAGAGAGAAAGAGCAGCTGAG 600
QY 907 GlnValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSer 926
Db 601 CAAGTTAGGAATCAACATTTGAATCCCAATGCCAAGAGGTTCAACCCAGCTTCTCTCT 660
QY 927 GlnProLysProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSer 946
Db 661 CAGCCAAAGCCTTCTACTCCCAACTTCACTCGGCCCTCAACACCAACCTAGCCCATCT 720
QY 947 MetValGlyHisGlnGlnProThrProValTyrThrGlnProValCysPheAlaProAsn 966
Db 721 ATGGTGGTCACTCAACAGCCAACTCCAGTTTATACCTCAGCTGTGTGTTTGCAACANAT 780
QY 967 MetMetTyrProValProValSerProGlyValGlnProLeuTyrProLleProMetThr 986
Db 781 ATGATGTATCCAGTCCAGTGAGCCCGGCGTGCAACCTTTATATACCAATACCTATGACG 840
QY 987 ProMetProValGlnGlnAlaLysThrTyrArgAlaValProAsnMetProGlnGlnArg 1006
Db 841 CCCATGCCAGTGAATCAGCCCAAGACATATAGAGCAGTACCAATATGCCCCCACAGCGG 900
QY 1007 GlnAspGlnHis 1010
Db 901 CAAGAACAGCAT 912

RESULT 5
BX375946 938 bp mRNA linear EST 26-APR-2004
LOCUS BX375946 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC020YG10 5-PRIME, mRNA sequence.
ACCESSION BX375946
VERSION BX375946.2 GI:46573234
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 938)
Li.W.B., Gruber.C., Jessee,J. and Polayes.D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi.30448458.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

```

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6451.r

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?e=CS0DC020BD05QPl&c=6451.r>.

Location/Qualifiers

1..938

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DC020YG10"

/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

#### FEATURES

source

#### Alignment Scores:

Pred. No.: 2,87e-70 Length: 938

Score: 1530.00 Matches: 301

Percent Similarity: 96.78% Conservative: 0

Best Local Similarity: 96.78% Mismatches: 8

Query Match: 21.98% Indels: 2

DB: 5 Gaps: 0

US-10-802-228-2 (1-1312) x BX375946 (1-938)

```

QY 690 ArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyProValLeuAlaSerPro 709
Db 3 CGGATCCAGACANACAGTATTGGAATATCCCCAGTGGGCCAGTCTTGTCTTCTCCC 62
QY 710 GlnAlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAlaSerProThr 729
Db 63 CAAGCTGCTATTATTCCAACTGAAGCTGTGCCATGCCTATTCCAGCTGCATCTCTACG 122
QY 730 ProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLysAspSerArg 749
Db 123 CTTGCTAGTCTCGCATCGAACAGAGCTGTACCCCTTCTAGTGAGGCTAAAGATTCCAGG 182
QY 750 LeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnLysProAsnGlu 769
Db 183 CTTCAAGATCAGAGGAGAACTCTCTCAGGGAATAAGAAATATTAAACCAATGAA 242
QY 770 ThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerProValValSerGluHis 789
Db 243 ACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAGTTGTTCTGAACAT 302
QY 790 ArgLysGln-IleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSe 809
Db 303 AGAAACACAGCATTTGATGATTTAAAGAAATTTAAAGAAATGATTTTAGGTTACAGCCAAGTTC 362
QY 809 rThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAs 829
Db 363 TACTTCTGAATCTATGGATCACTACTTAAACAAAAATAGAGAGGAGAAAAATCAAGAGA 422
QY 829 pLeuIleLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSerSerSe 849
Db 423 TTTGATCAAGACAAAAATTTGAACCAAGTGCTAAGGATTTCTTCATTTCAAAATAGCAGCAG 482
QY 849 rAsnCysThrSerGlySerSerLysProAsnSerProSerProSerIleSerProSerIleLeuSe 869
Db 483 CAACTGTACCAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTTCAATACTTAG 542
QY 869 rAsnThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerPr 889
Db 543 TAACAGGAGCAACAAGGGGACCTCAGGTGCTCTTCCCAAGGGTTTCAGACTTCCAGCCC 602
QY 889 oAlaCysLysGlnGluLysAspLysGluGluLysLysAspAlaLysValVal 909
Db 603 AGCATGTAAACAGNAGNAGNCGNTAAGGNAGAGAGAAAGAGCCAGCTCAGCAAGTTAG 662

```

QY 909 gtysserThrIeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProIly 929  
 Db 663 GAAATCAACATTGAATCCCAATGCAAGAGAGTCAACCCACGTTCTCTTCAGCCAAA 722  
 QY 929 sProSerThrProThrSerProArgProGlnAlaGlnProSerProSerMetValG1 949  
 Db 723 GCCTTCTACTACCCCAACTTCACTCGGCTCAAGCAACACTAGCCCATCTATGGTGGG 782  
 QY 949 yHisGlnGlnProThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTy 969  
 Db 783 TCATCAACAGCCCAAMTCCAGTTTATCTACCTCAGCTCTTTGTTTGCACCAATATGATGA 842  
 QY 969 rProValProValSerProGlyValGlnProLeuTyrProIle-ProMetThrProMetP 989  
 Db 843 TWCAGTCCAGTGAGCCCGGCGTCAACCTTTATACCCMATTTACCTATGAGCCCGCATGC 902  
 QY 989 roValAsnGlnAlaLysThrTyrArgAla 998  
 Db 903 CAGTGAATCAAGCCCAAGACATATAGCA 931

RESULT 6  
 BP159935 876 bp mRNA linear EST 30-DEC-2003  
 LOCUS BX402110 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 DEFINITION BX402110 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 ACCESSION BP159935  
 VERSION BP159935.1 GI:40409408  
 KEYWORDS EST.  
 SOURCE Sus scrofa (pig)  
 ORGANISM Sus scrofa

REFERENCE  
 AUTHORS Uenishi, H., Echuchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,  
 Okumura, N., Hamagima, N. and Awata, T.  
 TITLE PEDE (Fig EST Data Explorer): construction of a database for ESTs  
 derived from porcine full-length cDNA libraries  
 JOURNAL Nucleic Acids Res. 32 (1), D484-D488 (2004)  
 PUBMED 14681463

COMMENT  
 Contact: Hirohide Uenishi  
 Animal Genome Laboratory, Genome Research Department  
 National Institute of Agrobiological Sciences  
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan  
 Tel: +81-29-838-8627  
 Fax: +81-29-838-8627  
 Email: huenishi@affrc.go.jp  
 EST project with full-length enriched cDNA libraries carried out in  
 Animal Genome Research Program (Japan) by National Institute of  
 Agrobiological Sciences and STAFF-Institute  
 Single pass sequencing of clones derived from oligo-capped cDNA  
 library  
 Vector sequences were eliminated by RepeatMasker version 2002/07/13  
 and crossmatch version 0.990319  
 Low quality bases were trimmed based on the quality values.

FEATURES  
 source  
 Location/Qualifiers  
 1..876  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9823"  
 /clone="THY010091A06"  
 /tissue\_type="thymus"  
 /dev\_stage="adult"  
 /clone\_lib="full-length enriched swine cDNA library, adult  
 thymus"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,79e-70 Length: 876  
 Score: 1525.00 Matches: 283  
 Percent Similarity: 92.26% Conservative: 3  
 Best Local Similarity: 91.23% Mismatches: 4  
 Query Match: 21.91% Indels: 20

DB: 3 Gaps: 2  
 US-10-802-228-2 (1-1312) x BP159935 (1-876)  
 QY 991 AsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAsp 1008  
 Db 1 AATCAAGCAAGACATATAGAGCAGGTAAGTACCAATATATACCCCAACACGCGCAAGAC 60  
 QY 1009 GlnHisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAla 1028  
 Db 61 CAGCATCATCAGAGCACCACCATGATGCACCCAGCCTCAGCAGCAGGTCCACCAATTGTTGCC 120  
 QY 1029 ThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGln 1048  
 Db 121 ACCCCAGCAGCTTATCCACACAATATGTTGCCTATATGCTCTCAGCAGTTCCTCAAAATCAA 180  
 QY 1049 ProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProVal 1068  
 Db 181 CCTCTTGTTCAGCATGTGCCACATTATCAGTCTCAGCATCTCTATGCTATAGTCCTGTA 240  
 QY 1069 IleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSer 1088  
 Db 241 ATACAGGGTAATGCTAGAAATGATGGCACCACCAACACATGCCAGCCTGGTTTATGATCT 300  
 QY 1089 SerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLys 1108  
 Db 301 TCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACACATGCGATGAT----- 348  
 QY 1109 LeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeu 1128  
 Db 349 -----GTTTCCACTGGCTTCCCTT 366  
 QY 1129 AlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnPro 1148  
 Db 367 GCTCAACAGTATGCGCACCCTAATGCTACCTGACCCACACATATCCACATCTCCTCAGCCT 426  
 QY 1149 SerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaProSer 1168  
 Db 427 TCAGCTACTCTCCACTGGACAGCAGCAACCCAGCATGGTGGAGGCCATCTCTGCACCCAGT 486  
 QY 1169 ProValGlnHisHisGlnHisGlnAlaGlnAlaLeuHisLeuAlaSerProGlnGln 1188  
 Db 487 CCTGTTCAGACACCATCAGCACCAGGCTGCCAGGCTCTCCATCTGGCCAGTCCACACAG 546  
 QY 1189 GlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSer 1208  
 Db 547 CAGTCAGCCATTACCATGAGGCTCGTCCAAACCCACCTTCCATGACACCTGCCTCC 606  
 QY 1209 AsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHis 1228  
 Db 607 AATACGCAGTCGCCACAGAATAGCTTCCCAACAGCACAACAGACTGTCTTCACGATCCAT 666  
 QY 1229 ProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisValProGlnAla 1248  
 Db 667 CCTTCTCATGTTTCAGCCGGCATACCAATCCACCCACATGGCCCATGTACCTCAGCCT 726  
 QY 1249 HisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMet 1268  
 Db 727 CATGTACAGTCAGGAATGGTTCTTCTCATCTCAACTGCCCATGGCCCAATGATGCTAATG 786  
 QY 1269 ThrThrGlnProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProIle 1288  
 Db 787 ACGACACAGCCACCCGGCGTCCCGCAGCGCCCTCGCTCANAGTGCACTACAGCCCAT 846  
 QY 1289 ProValSerThrThrAlaHisPheProTyr 1298  
 Db 847 CCAGTCTCGAACACAGCGCATTTCCCTTAT 876

RESULT 7  
 BX402110  
 LOCUS BX402110 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 DEFINITION BX402110 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 CDNA clone CS0DC015YB14 5-PRIME, mRNA sequence.







EcORI/XhoI-digested pCS107. Reference for library construction: Current Genomics 4, 635-644. Library constructed by Michelle Tabb and Bruce Blumberg (Dept of Developmental and Cell Biology, University of California, Irvine)."

# ORIGIN

Alignment Scores:  
Pred. No.: 5,24e-67 Length: 925  
Score: 1467.50 Matches: 272  
Percent Similarity: 86.20% Conservative: 9  
Best Local Similarity: 83.44% Mismatches: 24  
Query Match: 21.08% Indels: 21  
DB: 8 Gaps: 3

US-10-802-228-2 (1-1312) x DR838409 (1-925)

QY	934	ProThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnPro	953
DB	3	CCAACTCAGCTCGTCCCAAACTCAGCAAGTCCATCATGTTGGATCATGACGACCA	62
QY	954	ThrProValThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal	973
DB	63	ACGCTGTGTATACCCAGCAGTATGTTTGACCCGAATATGATGATCCAGTCTCTGTG	122
QY	974	SerProGlyValGlnProLeuTyrProLeProMetThrProMetProValAlaGlnAla	993
DB	123	AGCCTGGAGTTCAGCCATTATCTCCATTCATGACTACATGAGTCCAGTGAACAGGCG	182
QY	994	LyethTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAspGlnHis	1011
DB	183	AGACTTACAGACAGAGTAAAGTACCAATATGCCCGCAACGGCAAGATCAGCACAC	242
QY	1012	GlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProLeAlaAlaThrProPro	1031
DB	243	CAGAAATCAGATGATCATCAGTATCTGACGAGGACCAATAGTGTGTACCCACCT	302
QY	1032	AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGlnProLeuVal	1051
DB	303	GCCTATTCTGCCAGTATGCTGCTTACAGCCCTCAACAGTTTCTTAATCAACCACTATG	362
QY	1052	GlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGly	1071
DB	363	CAGCATGTACAGCACTATCATGTCACAGCACCCATGTATATATAGTCTGTGATACAGGG	422
QY	1072	AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAla	1091
DB	423	AACACAGAATGATGGCACCACCATCTCATGACAGCTGGCTTGTATCTCTCGCT	482
QY	1092	ThrGlnTyrGlyAlaHisGlnGlnThrHisAlaMetTyrAlaCysProLysLeuProTyr	1111
DB	483	GCACAGTATGCCACTCTCGAACAACCTCATACCATGTAT-----	521
QY	1112	AsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGln	1131
DB	522	-----GTTTCAACTAGCTCATTAGCACAGCAG	548
QY	1132	TyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThr	1151
DB	549	TATGCTCACCACAAATGAGCGCTGACCCACCCATCTCATCTCTCATGCTTACGCAACT	608
QY	1152	ProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaProSerProValGln	1171
DB	609	CCTACCGGCCACAGCAAGTCAACATGGGGGAAGTCAACCCAGTCTCAAGTCCAGTTTCA	668
QY	1172	HisHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGlnGlnSerAla	1191
DB	669	CACCATCAACATCAGGCATCACAAGCTCTGCATCTGGCTAAACCAACAGCAGCTGTCA	728
QY	1192	IleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSerAsnThrGln	1211
DB	729	ATTTTACCATGACGAGCTTGCTTCCAAACGCCCTGCCATGACACCAAGCTCCACGCA	788

QY	1212	SerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHis	1231
DB	789	TCCCAACAGAGAGCTTCCCA---ACTCAACAGAGCTGTGTTTACCATTACCCATCACAT	845
QY	1232	ValGlnProAlaTyrThrAsnProProHisMetAlaHisValProGlnAlaHisValGln	1251
DB	846	GTCCAAGCAGCATATACCAATCCGCTCACATGGCCCATGTGCAACAGGCACATGTACAG	905
QY	1252	SerGlyMetValProSer	1257
DB	906	TCTGGAATGGTTCCTTCT	923
RESULT	10		
LOCUS	BQ428363		
DEFINITION	BQ428363	877 bp mRNA linear EST 24-MAY-2002	
ACCESSION	BQ428363	AGENCOURT_7858592 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6175163	
VERSION	BQ428363.1	5', mRNA sequence.	
KEYWORDS	EST.	GI:21167439	
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 877)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) CDNA library Arrayed by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM13549 row: h column: 12 High quality sequence stop: 652.		
FEATURES	Location/Qualifiers		
-source	1..877 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6175163" /tissue_type="leiomyosarcoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_71" /notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.01e-66	Length:	877
Score:	1461.50	Matches:	279
Percent Similarity:	91.56%	Conservative:	3
Best Local Similarity:	90.58%	Mismatches:	4
Query Match:	21.00%	Indels:	22
DB:	5	Gaps:	2
US-10-802-228-2 (1-1312) x BQ428363 (1-877)			
QY	880	ThrSerGlnGlyValGlnThrSerProAlaCysLysGlnGluLysAspLysGlu	899
DB	1	ACTTCCCAAGGGTTCAGACTTCCAGCCAGCATGTAACCAAGAAAGACGATAAGAA	60
QY	900	GluLysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGlu	919
DB	61	GAGAGAAAGACGAGCTGAGCAAGTTAGGAATCAACATTGAATCCCATGCAAGAG	120
QY	920	PheAsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgPro	939

```
Db 121 TTCAACCCACGGTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCT 180
Qy 940 GlnAlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGln 959
Db 181 CAAGCACACCTAGCCCACTATGTGGTGCATCAACAGCCCACTCCAGTTTATCTAG 240
Qy 960 ProValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnPro 979
Db 241 CCTGTTGTGTTTGCACAAATATGATGATTCAGTCCAGTCCAGCCAGCCGCGTCAACT 300
Qy 980 LeuTyrProIleProMetThrProMetProValAsnGlnAlaValThrTyrArgAlaVal 999
Db 301 TTATACCCCAATACCTATGAGCCCAAGCCAGTGAATCAAGCCCAAGACATATAGAGAGTA 360
Qy 1000 ProAsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAla 1019
Db 361 CCAATATGCCCAACAGCGGCAAGACCAAGCATCATCAGATGCCATGATGACCCAGCG 420
Qy 1020 SerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAla 1039
Db 421 TCAGCAGCGGGGCCCAACCGATGTCAGCCACCCACCACTTACTCCAGGCAATATGTTGCC 480
Qy 1040 TyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSer 1059
Db 481 TACAGTCCTCAGCAGGTCCCAATCAGCCCTTGTTCAGCATGTGCCAATATCAGTCT 540
Qy 1060 GlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProPro 1079
Db 541 CAGCATCCTCATGTCTATAGTCTGTATACAGGGAATGCTAGATGATGATGACCAACCA 600
Qy 1080 ThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGln 1099
Db 601 ACACAGCCCGCAGCTGTTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAG 660
Qy 1100 ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPhe 1119
Db 661 ACGATGCGCATGTAT----- 675
Qy 1120 TyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeu 1139
Db 676 -----GTTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTGT 726
Qy 1140 HisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGln 1159
Db 727 CACCCACATATCCACACCCTCAGCCTTTCAGCTACCCCACTGGCAGCAGCAAGCCAA 786
Qy 1160 HisGlyGlySerHisPro-AlaProSer-ProValGlnHis-HisGlnHisGlnAla--- 1177
Db 787 AATGTGGGAAGTCACTCTGGCACCCAGTCCCTGGTTCAGCACCCCATCAGCACCGGGCGG 846
Qy 1178 AlaGlnAlaLeuHisLeu 1183
Db 847 CCCAGGGTCTCCATCTG 864

RESULT 11
BQ716397
LOCUS
DEFINITION
AGENCY: 8477984 Lupski sympathetic_trunk Homo sapiens cDNA clone
IMAGE: 6196657 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 859)
NIH-MGC http://mgi.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
```

Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13605 row: h column: 02  
High quality sequence stop: 659.

FEATURES  
Location/Qualifiers  
1..859  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6196657"  
/sex="male"  
/tissue\_type="sympathetic trunk"  
/dev\_stage="adult, 16 yr"  
/lab\_host="DH10B"  
/clone\_lib="Lupski sympathetic trunk"  
/note="Vector: pCMV-Sport6 (Life Technologies); Site\_1: Not1; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCTCCG-3' and 5'-GACTAGTTCTAGATCGCAGGGCGGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

ORIGIN  
Alignment Scores:  
Pred. No.: 4,68e-66 Length: 859  
Score: 1448.50 Matches: 277  
Percent Similarity: 98.95% Conservatives: 5  
Best Local Similarity: 97.19% Mismatches: 2  
Query Match: 20.81% Indels: 1  
DB: 5 Gaps: 1  
US-10-802-228-2 (1-1312) x BQ716397 (1-859)

Qy 555 CysProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuPro 574  
Db 3 TGCCCATCTCTCTCTCTGCGCCACCTTCTCGTACCACTCAGGTCCCACTCTTCCA 62  
Qy 575 ProArgAlaAlaThrProThrArgProProSerArgProProSerArgProSerArgPro 594  
Db 63 CTTGGGCGAGCAGCCCTTACACGCGCGCCCTCCAGGCCCTCCGCGCCATCCAGACCC 122  
Qy 595 ProSerHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMet 614  
Db 123 CCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCTGTCTACTATGCCCTAAACGCATG 182  
Qy 615 SerSerGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisArg 634  
Db 183 TCTTCAGAGGCGCTCCCAAGGATGTCCCAAGGCCAGCCAGCATCTCTCGAAATCACAGA 242  
Qy 635 ValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProPro 654  
Db 243 GTTTCTGCTGGAGGGTTCCATATCCAGTGGCTAGAAATTTGTATCCCAACCCACCC 302  
Qy 655 SerGluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerSer 674  
Db 303 AGTGAAGCAGCTACTCTCTCCAGTAGCAAGGACCACTCCCTCGGGGGAACTGGTCACTCA 362  
Qy 675 ValValSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGln 694  
Db 363 GTGGTCAGTGGGGTTCCAAAGATTATCCCTAAAACCTCATAGACCAGGCTCTCCAGACAG 422  
Qy 695 AsnSerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleIle 714

Db 423 AACAGTATTGGAAATACCCAGTGGCCAGTCTTCTGCTTCTCCCAAGCTGGTATTATT 482

Qy 715 ProThrGluAlaValAlaMetProIleProAlaAAlaSerProThrProAlaSerProAla 734

Db 483 CCAACTGAAGCTGTGTGCCATGCTATTCCAGCTGCATCTCCTACGCTGCTAGTCTCGCA 542

Qy 735 SerAsnArgAlaValThrProSerGluAlaLysAspSerArgLeuGlnAspGlnArg 754

Db 543 TCGAACAGAGCTGTGTACCCCTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGG 602

Qy 755 GlnAsnSerProAlaGlyAsnLysGluAsnLysGluAlaLysProAsnGluThrSerProSerPhe 774

Db 603 CAGAACTCTCTCGCAGGAAATAAGAAATATTAAACCCCAATGAACATCACTAGCTTC 662

Qy 775 SerLysAlaGlnAsnLysGlyLysSerProValValSerGluHisArgLysGlnLysAsp 794

Db 663 TCAAAAGCTGMAAACCAAGATATATCACCAGTTGTTCTGAACATAGAAAAACAGATTGAT 722

Qy 795 AspLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMet 814

Db 723 GATTTAAGAAATTTAAGATGATTTTAGGGTACAGCCAAAGTTCTACTTCTGAATCTATG 782

Qy 815 AspGlnLeuLysAsnLysAsnArgGluGlyLysSerArgAspLeu---LysLysAsp 833

Db 783 GATCACTACTAAACAAATAAGAGAGGGGAGAAAAATCCAGAGATTTGGATCAAGAC 842

Qy 834 LysLysGluProSer 838

Db 843 CAAATGAACCAAGT 857

RESULT 12

CA489164

LOCUS

DEFINITION AGENCOURT\_10853363 MAPcL Homo sapiens cDNA clone IMAGE:6721472 5', mRNA sequence.

ACCESSION CA489164

VERSION

KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini, Homnidae; Homo.

1 (bases 1 to 889)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-rcmail.nih.gov

Tissue Procurement: Kristi A. Eglund, Ira Pastan

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

plate: MAM14282 row: k column: 08

High quality sequence stop: 683.

Location/Qualifiers

1. 889

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6721472"

/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCaP"

/lab\_host="EMDH108"

/clone\_lib="WAPcL"

/note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dr. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg,

Bungkook Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,01e-65 Length: 889

Score: 1442.50 Matches: 284

Percent Similarity: 96.28% Conservative: 1

Best Local Similarity: 95.95% Mismatches: 8

Query Match: 20.72% Indels: 4

DB: 6 Gaps: 2

US-10-802-228-2 (1-1312) x CA489164 (1-889)

Qy 738 AlaValThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSer 757

Db 2 GCTGTTACCCCTTCTAGTGGCTAAAGATTCCAGGCTTCAAGATCAGAGGAGAACTCT 61

Qy 758 ProAlaGlyAsnLysGluAsnLysProAsnGluThrSerProSerPheSerLysAla 777

Db 62 CCTGCGAGGGAATAAGAAATATTAAACCCCAATGAACATCACTAGCTTCTCAAAAGCT 121

Qy 778 GlnAsnLysGlyLysSerProValValSerGluHisArgLysGlnLysAspLys 797

Db 122 GAAACAAAGGTATATCACCAGTTGTTCTGAACATAGAAAAACAGATTCATGATTTAAAG 181

Qy 798 LysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeu 817

Db 182 AAATTTAAGAATGATTTTAGGTTACAGCCAAAGTTCTACTTCTCAATCTATGATCAACTA 241

Qy 818 LeuAsnLysAsnArgGluGlyLysSerArgAspLeuLysAspLysLysGluPro 837

Db 242 CTAAACAAATAATAGAGAGGAGAAAAATCAAGAGATTTCATCAAGACACAAAATTTGAACCA 301

Qy 838 SerAlaLysAspSerPheLysGluAsnSerSerSerAsnCysThrSerGlySerSerLys 857

Db 302 AGTGCTAAGGATCTTTCATTGAAATATAGCAGCAACTGTACCACTGGCAGCAGCAAG 361

Qy 858 ProAsnSerProSerLysSerProSerLysLeuSerAsnThrGluHisLysArgGlyPro 877

Db 362 CCGAATAGCCCCAGCAGCTTTCCCTTCAATCTAGTAAACAGCAGCACAAGAGGGGACCT 421

Qy 878 GluValThrSerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAsp 897

Db 422 GAGGTCACTTCCCAAGGGGTTCCAGACTTCCAGGCCAGCATGTAAACAAGAAAGACGAT 481

Qy 898 LysGluGluLysLysAspAlaGluGlnValArgLysSerThrLeuAsnProAsnAla 917

Db 482 AAGGAAGAGAAAGAAAGCAGCTGAGCAAGTTAGGAATTCACATTTGAATCCCAATGCA 541

Qy 918 LysGluPheAsnProArgSerPheSerGlnProLysProSerThrThrProThrSerPro 937

Db 542 AAGGAGTTCAACCCACGTTCTTCTCTCAGCCAAAGCTTCTACTACCCCACTTCACT 601

Qy 938 ArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyr 957

Db 602 CGGCCTCAAGCACAACCTAGCCCATCTATGGTGGGTCTATCAACAGCAGCAACTCCAGTTAT 661

Qy 958 ThrGlnProValCysPheAlaProAsnMetMetTyrProValProValSerProGlyVal 977

Db 662 ACTCAGCCCTGTTGTTTTCACCAAAATATGATGATCCAGTCCAGTCCAGCCAGCGCTG 721

Qy 978 GlnProLeuTyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArg 997

Db 722 CAACCTTTATA-CCAATACCTATGACGCCCATCCAGTGAATCAAGCCAAGACATATAGA 780

Qy 998 Ala-----ValProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMet 1015

Db 781 GCAGGTAAAGTACCAAAATATGCCCCACAGCGNCAAGACCAGCATCATCCAAAGGCCATG 840

Qy 1016 MetHisProAlaSerAlaAlaGlyPro---ProIleAlaAlaThrPro 1030

|||||

Db 841 ATGACCCAGCGCTACAGCGGGGGCCACCGCAATTGCAGCCACCCCA 888

RESULT 13  
 DR867406 917 bp mRNA linear EST 29-JUL-2005  
 LOCUS JGI\_CABG9511.fwd NIH XGC tropStol Xenopus tropicalis cDNA clone  
 DEFINITION IMAGE:7840708 5', mRNA sequence.  
 ACCESSION DR867406.1 GI:71523996  
 VERSION DR867406  
 KEYWORDS EST.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 1 (bases 1 to 917)  
 Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,  
 Brokstein,P. and Lindquist,E.A.  
 DOE Joint Genome Institute Xenopus tropicalis EST project  
 Unpublished (2004)  
 Other ESTs: JGI\_CABG9511.rev  
 Contact: Lindquist,E.A., Richardson,P.  
 DOE Joint Genome Institute  
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 Tel: 925 296 5600  
 Fax: 925 296 5710  
 Email: cdna@jgi-psf.org  
 Tissue Procurement: Robert M. Grainger  
 cDNA Library Preparation: Bruce Blumberg Laboratory, University of  
 California, Irvine  
 DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov  
 Clone Distribution: I.M.A.G.E. Consortium/LNL:  
 http://image.llnl.gov  
 Naming Conventions: EST name is generated by the concatenation of  
 the JGI Clone Id and the direction of sequencing. The suffix '.fwd'  
 indicates a forward sequencing read of the insert. It does not  
 necessarily reflect the orientation of the insert.  
 Plate: CABG 0097 row: n column: 2  
 High quality sequence stop: 814.  
 Location/Qualifiers  
 1..917  
 /organism="Xenopus tropicalis"  
 /mol\_type="mRNA"  
 /strain="N6 (Nigerian 6th generation inbred)"  
 /db\_xref="taxon:8364"  
 /clone="IMAGE:7840708"  
 /tissue\_type="Stomach"  
 /dev\_stage="Adult"  
 /lab\_host="ElectroMAX DH10B T1 Phage Resistant cells"  
 /clone\_lib="NIH XGC tropStol"  
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The  
 library was prepared from 5'ug of poly A+ RNA by oligo-dT  
 priming  
 (5'-ACTAGTCGGCGCGCTAGGCTCGAGTTTCTTTTCTTTTCTTTT-3') and  
 Stratascript reverse transcriptase. After ligation of  
 EcoRI adapters (5'-AATTCGGCAGGAG-3') followed by kinasing  
 adapters and by XhoI digestion, the cDNA was size selected  
 by chromatography on Sepharose CL-2B columns and fractions  
 containing cDNAs larger than 1000 bp were ligated into  
 EcoRI/XhoI-digested pCS107. Reference for library  
 construction: Current Genomics 4, 635-644. Library  
 constructed by Michelle Tabb and Bruce Blumberg (Dept of  
 Developmental and Cell Biology, University of California,  
 Irvine)."

ORIGIN

Alignment Scores:  
 Pred. No.: 4,46e-65 Length: 917  
 Score: 1430.50 Matches: 269  
 Percent Similarity: 85.23% Conservative: 8  
 Best Local Similarity: 82.77% Mismatches: 27  
 Query Match: 20.55% Indels: 21  
 DB: 8 Gaps: 4

FEATURES  
 source

US-10-802-228-2 (1-1312) x DR867406 (1-917)

QY 983 IleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValProAsnMet 1002  
 DB 5 ATTTCAATGACTACAAATGCCAGTGAACAGGCCAAGACTTACAGAGCGATACCCNATATG 64  
 QY 1003 ProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSerAlaAla 1022  
 DB 65 CCCAGCAACGGCAAGATCAGCACCCAGCAATACATGATGATGATGATGATGATGATGATG 124  
 QY 1023 GlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyrSerPro 1042  
 DB 125 GGACCAACCAATAGTTGTTGATCCCACTTATGAGGATGATGATGATGATGATGATGATG 184  
 QY 1043 GlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGlnHisPro 1062  
 DB 185 CAACAGTTTCTTAATCAACCACTTATGAGGATGATGAGGATGATGAGGATGATGAGGATG 244  
 QY 1063 HisValTyrSerProValIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 1082  
 DB 245 CATGTATATAGTCTCTGTGATACAGGGGAACACAAAGATGATGGCCACCATCTCATGCA 304  
 QY 1083 GlnProGlnLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAla 1102  
 DB 305 CAGGCTGGCTTGTATCTCTCTCGCTGCACAGTATGCCATCTCTCAACAACTCATACC 364  
 QY 1103 MetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAla 1122  
 DB 365 ATGTAT----- 370  
 QY 1123 IleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHis 1142  
 DB 371 GTTTCAACTAGTCTATTAGCACACAGCATGTCTCCCAATGATGGCCCTGCACCCACAC 430  
 QY 1143 ThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGlyGly 1162  
 DB 431 CCAGCTCATCTCTCAGCCTTCCAGCAACTCTACCGGCCAACAGCAAAAGTCAACATGGGGA 490  
 QY 1163 SerHisProAlaProSerProValGlnHisGlnHisGlnAlaGlnAlaLeuHis 1182  
 DB 491 AGTCACCCAGCTCCAGTCCAGTTTCAGCACCATCAACATCAGGCATCAAGCTCTGCAT 550  
 QY 1183 LeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProPro 1202  
 DB 551 CTGGCTAACCAACAGCAGCAGTCTGCAATTTACCATGCAGGACTTGCTCCCAACGCCCT 610  
 QY 1203 SerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGln 1222  
 DB 611 GCCATGACACCCAGGCTCCACGCACAGTCCCAACAAAGCAGCTTCCCA...ACTCAACAG 667  
 QY 1223 ThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProHisMet 1242  
 DB 668 ACTGTGTTTACATTTACCCCATCATGTCTCAACGAGCATATACCAATCCGCCCTCACATG 727  
 QY 1243 AlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThrAlaHis 1262  
 DB 728 GCCCATGTGCAACAGGCACATGTACAGTCTGGAATGGTTCTCTCTCATCCAACCTGCTCAT 787  
 QY 1263 AlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeuAlaGln 1282  
 DB 788 ---CCCATGATGTTGATGACTGTCAACCACTCGGGGGGCTCAAGCTGCTTTGCTCNA 844  
 QY 1283 SerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThrHisPro 1302  
 DB 845 AGTGCCTTACAGCCCATTCCTGTC---TNCAGTCACACTTTTCTTATATGACATCCT 901  
 QY 1303 SerValGlnAlaHis 1307  
 DB 902 CCAGTGAAGCAGCAC 916  
 RESULT 14  
 CD616048/c

```

LOCUS      CD616048      830 bp      mRNA      linear      EST 12-JAN-2004
DEFINITION 56069370H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD616048
VERSION    CD616048.1 GI:40264312
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 830)
AUTHORS   Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE     Circular rapid amplification of cDNA ends for high-throughput
PUBMED    Genomics 84 (1), 205-210 (2004)
COMMENT   15203218
            Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES   Location/Qualifiers
            source
            1..830
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_lib="FLP"
            /note="Vector: pDrive Cloning Vector"
ORIGIN
Alignment Scores:
Pred. No.:      4,62e-64      Length:      830
Score:          1410.00      Matches:     264
Percent Similarity: 97.06%      Conservative: 0
Best Local Similarity: 97.06%      Mismatches:  7
Query Match:     20.26%      Indels:      2
DB:              6          Gaps:         0
US-10-802-228-2 (1-1312) x CD616048 (1-830)
QY      1040 TyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSer 1059
DB      829 TACAGTCCTCAGCAGTTCCTCCAAATACCCCTGTTTCAGCATGTGCCACATATCATGCT 770
QY      1060 GlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProPro 1079
DB      769 CAGCATCTCATGCTATAGTCTGTAATACAGGGTAATGCTAGATGATGGCACCACCA 710
QY      1080 ThrHisAlaGln-ProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluG 1099
DB      709 ACACAGCCCGCCCTGGTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCA 650
QY      1099 nThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPh 1119
DB      649 GACGCATCGCATGTATGATGTCTCCAA-ATTACCATACAACAGGAGACAAGCCCTCTCTT 591
QY      1119 eTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLe 1139
DB      590 CTACTTTGCCATTCTCCACGGCTCCCTTGCTCAGCAGTATGGCAGCCCTAACCGCTACCT 531
QY      1139 uHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerG 1159
DB      530 GCACCCACATACCTCCACACCTCAGCTTCCAGTACCCCTGAGCAGCAGCAAGGCCA 471
QY      1159 nHisGlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaG 1179
DB      470 ACATGGTGGAGTTCATCTGCACCCAGTCTCTGTTTCAGCACCGTCAGCAGCCGCGCCA 411
QY      1179 nAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaPr 1199
DB      410 GGCTCTCATCTGGCCAGTCCACAGCAGCAGTACGCCATTATACCCGGGGCTTGCGCC 351
QY      1199 oThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAl 1219

```

```

DB      350 AACTCCACCTCCATGACACCTGCTCCAAACGCGAGTCGCCACAGTAAGTTTCCCGCAGC 291
QY      1219 aAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPr 1239
DB      290 AGCACAAACAGACTGCTTTTACGATCCATCTCTCAGCTTCAGCCGGGTATACCAACCC 231
QY      1239 oProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisPr 1259
DB      230 ACCCCACATGGCCACGTCACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCATCC 171
QY      1259 oThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAl 1279
DB      170 AACTGCCCATGCGCCAAATGATGCTTAATGACACACAGCACCCCGGGTCCCGAGCGCC 111
QY      1279 aLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMe 1299
DB      110 CCTCGCTCAAAGTGCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTATAT 51
QY      1299 tThrHisProSerValGlnAlaHisGlnGln 1310
DB      50 GACGACCCCTTCAGTACAAAGCCACCCACCAACAG 17
RESULT 15
CX568338      865 bp      mRNA      linear      EST 12-JAN-2005
LOCUS      UI-M-180-cuv-e-17-0-UI.r1 NIH_BMAP_180 Mus musculus cDNA clone
DEFINITION IMAGE:30946480 5', mRNA sequence.
ACCESSION  CX568338
VERSION    CX568338.1 GI:57595367
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 865)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
            cDNA library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
Seq primer: pYX-5.
            Location/Qualifiers
            1..865
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE:30946480"
            /tissue_type="whole eye"
            /dev_stage="newborn (1,5,15 days)"
            /lab_host="DH10B (T1 phage resistant)"
            /clone_lib="NIH_BMAP_180"
            /note="Organ: Eye; Vector: pYX-Asc; Site:1: EcoR I;
            Site 2: Not I; The library was constructed according
            Bonafido, Lennon and Soares, Genome Research, 6:791-806,
            1996. Denatured RNA was size fractionated on a 1% agarose
            gel. First strand cDNA synthesis was primed with oligo-dT
            size selected according to mRNA size fraction, ligated
            with EcoR I adaptor, digested with NotI and then cloned
            directionally into pYX-Asc vector. The library tag
            sequence located between the Not I site and the polyA tail
            is AATAATTACG. This library was created for the University

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 23, 2005, 06:06:48 ; Search time 513 Seconds  
(without alignments)  
4546.123 Million cell updates/sec

Title: US-10-802-228-2  
Perfect score: 6961  
Sequence: 1 MRSAAAPRSPAVATESRRF.....TAHFFPYMTHPSVQAHHQOOL 1312

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues  
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl  
-Q/cgn2\_1/USPTO\_spool\_p/US10802228/runat\_21122005\_110225\_9729/app\_query\_fasta\_1.1479  
-DB=issued Patents NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10802228 @CGN 1.1 328 @runat\_21122005\_110225\_9729 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	3	US-09-041-886-18
2	6961	100.0	4481	3	US-09-648-281-1
3	6961	100.0	4481	3	US-09-707-919A-20
4	6961	100.0	4481	3	US-09-083-268-2
5	6961	100.0	4481	3	US-08-981-998A-2
6	6891.5	99.0	4484	3	US-09-949-016-4498
7	5606.5	80.5	3798	3	US-08-981-998A-4
8	1876	27.0	1257	3	US-09-648-281-11
9	1876	27.0	1257	3	US-09-083-268-4

10	1876	27.0	1257	3	US-08-981-998A-19	Sequence 19, Appl
11	1245.5	17.9	151088	3	US-09-949-016-16240	Sequence 16240, A
12	1012.5	14.5	623	3	US-09-043-303-5	Sequence 5, Appl
13	857	12.3	516	3	US-09-083-268-1	Sequence 1, Appl
14	857	12.3	516	3	US-08-981-998A-1	Sequence 1, Appl
15	787	11.3	459	3	US-09-707-919A-13	Sequence 13, Appl
16	643	9.2	355	3	US-09-043-303-1	Sequence 1, Appl
17	568	8.2	30690	3	US-09-914-286-1	Sequence 1, Appl
18	535.5	7.7	13987	2	US-08-804-227C-13	Sequence 13, Appl
19	535.5	7.7	44377	2	US-08-804-227C-7	Sequence 7, Appl
20	535.5	7.7	44377	2	US-08-804-198-1	Sequence 1, Appl
21	520.5	7.5	47677	2	US-09-949-002-668	Sequence 668, App
22	514	7.4	43280	2	US-08-804-227C-1	Sequence 1, Appl
23	501	7.2	34094	3	US-09-292-034-1	Sequence 1, Appl
24	498.5	7.2	4725	3	US-09-410-551B-24	Sequence 24, Appl
25	498.5	7.2	4725	3	US-09-940-316B-24	Sequence 24, Appl
26	497.5	7.1	33529	3	US-09-144-085-3	Sequence 3, Appl
27	497.5	7.1	50937	3	US-09-428-517-1	Sequence 1, Appl
28	494.5	7.1	8090	3	US-09-902-540-855	Sequence 855, App
29	493	7.1	31422	3	US-09-914-286-2	Sequence 2, Appl
30	490.5	7.0	80161	3	US-09-036-987A-1	Sequence 1, Appl
31	490.5	7.0	80161	3	US-09-370-700-1	Sequence 1, Appl
32	490.5	7.0	80161	3	US-09-603-207-1	Sequence 1, Appl
33	482	6.9	19598	2	US-09-902-540-1143	Sequence 1143, Ap
34	481.5	6.9	28958	2	US-08-258-261B-6	Sequence 6, Appl
35	481.5	6.9	28958	2	US-08-456-837-6	Sequence 6, Appl
36	481.5	6.9	28958	2	US-08-457-342-6	Sequence 6, Appl
37	481.5	6.9	28958	2	US-08-457-646A-6	Sequence 6, Appl
38	481.5	6.9	28958	2	US-08-458-076A-6	Sequence 6, Appl
39	481.5	6.9	28958	2	US-08-764-233A-4	Sequence 4, Appl
40	481.5	6.9	28958	2	US-08-457-335A-6	Sequence 6, Appl
41	481.5	6.9	28958	2	US-08-729-214-6	Sequence 6, Appl
42	481.5	6.9	28958	3	US-09-028-934-6	Sequence 6, Appl
43	481.5	6.9	47981	2	US-09-679-279-1	Sequence 1, Appl
44	481.5	6.9	49377	2	US-08-764-233A-1	Sequence 1, Appl
45	481	6.9	8438	2	US-07-945-283-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-041-886-18  
; Sequence 18, Application US/09041886  
; Patent No. 6235872  
; GENERAL INFORMATION:  
; APPLICANT: Bredesen, Dale E.  
; APPLICANT: Rabizadeh, Sharoz  
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence  
; TITLE OF INVENTION: Polypeptides and Methods of Use  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/041,886  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 2626  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949

! INFORMATION FOR SEQ ID NO: 18:

! SEQUENCE CHARACTERISTICS:

! LENGTH: 4481 base pairs

! TYPE: nucleic acid

! STRANDEDNESS: single

! TOPOLOGY: linear

! MOLECULE TYPE: DNA (genomic)

! FEATURE:

! NAME/KEY: CDS

! LOCATION: 163..4099

US-09-041-886-18

Alignment Scores:

Pred. No.: 1,33e-256 Length: 4481

Score: 6961.00 Matches: 1312

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-041-886-18 (1-4481)

```
Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGCGCTCAGCGCGCGCAGCTCTCGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTC 222
Qy 21 AlaAlaAlaArgTrrProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db 223 GCCGAGCAGGTGGCCCGGGTGGCGTCCGCTCCAGCGCGCGCGCGGAGCGGCGG 282
Qy 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 283 GCGCGCGTGGCGCGCGCGCGGACGATCCCTCCGCGCGCGCTCCCGCGCGCGCGCC 342
Qy 61 GlyProProSerArgGlnSerSerProProSerAlaAlaSerAspCysPheGlySerAsn 80
Db 343 GCGCGCGCTCCCTCCGCGAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 402
Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
Db 403 GGCACCGCGCGCGCGGTTTCGCGCGCGCTCCCGCGCGCTCTTGGTCTCGGCGGCGCT 462
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaAlaSerProGlyAlaProProAlaAla 120
Db 463 CCGCGCGCGCTTCGTCGTCTCTCCCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCG 522
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 523 CCNACCGCGCGCTCCCGCTCGCGCGCGCTCGCTCCCGCGCGCTCCCGCGCTCTCCTTG 582
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 583 GCGCGCGCGCTCCCGCTGTCGCGCGCGCTCGCGCGCGCTCGCGCGCTATGGCGCTCACC 642
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 643 ATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
Db 763 AGCGCGCTTCTAGCGTCCCGCGCGCGCGCTTCGCGCTTCGCTCTCGCTCTCCTCG 822
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
Db 823 TCCCTCGCGCAGCGCTCCCTCTCGGTGTCGCGCGAGCTCCGCGCGCGCGCGCGCGCGCG 882
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db 883 CTGGCAGAGGTGCGAACAAGTAGCAAGAGCTGCTCAGTCTACGATTTCTTTTATGGA 942
```

```
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db 943 ATCTATGCAAAATATGAGGATGGTTCATATACTTACATCAGTTGTTGGCTCCAAATGTCAA 1002
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 1003 GTACAAGTGAAAAATGGAGGTATATATGAAGGAGTTTTTAAACCTTACAGTCCGAAGTGT 1062
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db 1063 GATTTGGTACTTGTATGCGGCACATGAGAAAAGTAGACAGATCCAGTTCGGGGCCGAAACGT 1122
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValIleGlnPheLys 340
Db 1123 GAAGAAATAATGGAGAGTATTTTGTTCAAATGTTTTCAGACTTGTGTGTTAGTACAGTTTAAA 1182
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db 1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCCTAAA 1242
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380
Db 1243 GTGAATGGCGAACAACAAGAGAGGACCTGGAGGCCCTGGGATGCGAGGTGAACACTCACGCC 1302
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMet 400
Db 1303 AATCAGGAACCTTGAGGCTTTTGGAAAAATGACGTAATCTAATGGATGGGATCCCAATGATATG 1362
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db 1363 TTTTCGATATATGAAGAAAAATATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTCTCG 1422
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1423 TATACAGTGCCTTAGAAGAGATAACTCAGAGAAGATTTTTTAAACCGGAAGCAAGGCA 1482
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACCACTTAGCAGAAAGAAATTTAGTCAAGTGCCTCCAGTACAAAGCTCAGTGGCCTGGAA 1542
Qy 461 AsnAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATCATGATAGGAGTGAGGAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db 1603 GAGGGGCACAGCATAAACACTAGGGAAAAATAAATATATCTCTCTGGACAAAGAAATAGA 1662
Qy 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTGGGAAGTGGGAGACAGAAATTCACCGGTATGGCCACGCTGGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTCAACCCGAAATCTCTGGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCTCTCTCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCACCTTCTCGCTACCACTCAGGTCCCAACTCTCTTCCACCTCGGGCAGCCACCCCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACACGGCGCGCTCCAGCGCGCGCTCCAGACCCCGCTCTCACCCCTCTCTCTCTCTCTCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGGTCTCCAGCTCTCTGCTCTCTATGCTTAACGATGCTTTCAGAAAGGCGCTCCA 2022
```

QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
DB 2023 AGGATGTCCTCCAAAGGCCAGGACATCTCGAAATACACAGAGTTTCTGCTGGGAGGGT 2082  
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
DB 2083 TCCATATCAGTGGCCCTAGAAATTTGTATCCACACCCACCCAGTGAAGCAGCTACTCT 2142  
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTpsSerSerValValSerGlyValPro 680  
DB 2143 CCAGTAGCAAGACCAAGTCCCTCGGGGGAACGTGGTTCATCAGTGGTGGGTTCCA 2202  
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
DB 2203 AGATTATCCCTAAAACCTCATGACCCAGGTCTCCAGACAGAACTAGTATTGGAATACC 2262  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
DB 2263 CCAGTGGGCGAGTCTTGTCTTCTCCCAAGCTGGTATTATTCCACTGAGCTGTGGC 2322  
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
DB 2323 ATGCCTATTCCAGCTCACTCTACGCTCTGCTAGTCTGATCGAACAGAGCTGTACC 2382  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCGAGAACTCTCCTGCGGG 2442  
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
DB 2443 AATAAGAAATATTAACCCCAATGAACATCACTAGCTTCTCAAAAGCTGAAACAAA 2502  
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
DB 2503 GGTATATCACCACTGTTCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAG 2562  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
DB 2563 AATGATTTTAGGTTACAGCCAAAGTCTTACTTCTGAATCTATGATCAACTACTTAAACAAA 2622  
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
DB 2623 AATAGAGGAGGAGAAATCAAGAGATTGTATCAAGACAAAATTAAGCAAGTGCTAAG 2682  
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
DB 2683 GATCTTTTCAATGAAATAGCAGCAGCACTGTACCAGTGGCAGCAGCAGCCGATAGC 2742  
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
DB 2743 CCCAGCATTTCCCTCTCAATACTTAGTAACACGAGCACAAAGAGGGGACCTGAGGTCCT 2802  
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900  
DB 2803 TCCCAAGGGGTTCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGATAGGAAGAG 2862  
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
DB 2863 AAGAAGACGCGAGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTTC 2922  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrProArgProGln 940  
DB 2923 AACCCAGGTTCTTCTCAGCCAAAGCCCTTACTACCCCAACTTCCACCTCGGCCCTCAA 2982  
QY 941 AlaGlnProSerProSerMetValGlyHisGlnProThrProValThrThrGlnPro 960  
DB 2983 GCACAACTAGCCCATATATGGTGGGTATCAACAGACCAACTCCAGTTTATATCTCAGCCT 3042  
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
DB 3043 GTTTGTTTGCACCAATATATGATGTATCCAGTCCAGTGAGCCAGCGGCTGCAACCTTTA 3102  
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000

DB 3103 TACCCAAATACCTATGACGCCCATGCCAGTGAATCAAGCCAGACATATAGACGAGTACA 3162  
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020  
DB 3163 AATATGCCCAACACAGCGCAAGACCAGCATCATCAGAGTGCCCATGATGACCCAGCGTCA 3222  
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaThrSerThrGlnTyrValAlaThr 1040  
DB 3223 GCAGCGGGCCCAACCATTTGCAGCCACCCACAGCTTACTCCACCAATATATGTCCTAC 3282  
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
DB 3283 AGTCTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATATATCAGTCTCAG 3342  
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
DB 3343 CATCTCTCATGTCTATAGTCTCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3402  
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
DB 3403 CACGCCCAAGCCCTGGTTAGTATCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3462  
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
DB 3463 CATGGATGTATGTCATGTCCCAAAATTTACATCAACAAGGAGACAAAGCCCTTCTTTCTAC 3522  
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
DB 3523 TTTGCCATTTCCACGGGCTCCCTTCTCAGCAGTATGCGCACCTTAACCTACCCTGTCAC 3582  
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160  
DB 3583 CCACATACTCCACACCCCTCAGCTTTCAGTACCCCTCAGCAGCAGCAAGCCCAACAT 3642  
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180  
DB 3643 GGTGAAGTCACTCTGACCCAGTCTGTTCAGCACCCTCAGCAGCCAGCCGCCAGGCT 3702  
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTACAGCAATTTACCACGGGGCTTGCGCAACT 3762  
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
DB 3763 CCACCCCTCATGACACCTGCTCCCAACAGCAGTCCGCCACAGAAATAGTTTCCAGCAGCA 3822  
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaThrThrAsnProPro 1240  
DB 3823 CAACAGACTGTCTTTACGATCTCTCTCAGTTCAGCCGCGGTATACCAACCCACCC 3882  
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
DB 3883 CACATGGCCCACTGCTACCTCAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCAACT 3942  
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
DB 3943 GCCCATGCGCCCAATGATGTCTAATGACACACAGCCACCCGCGGTCCCGAGCGCGCCCTC 4002  
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
DB 4003 GTTCAAGTGCACCTACAGCCCATTCAGTCTCGAACAAAGCGCATTTCCCTATATGACG 4062  
QY 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312  
DB 4063 CACCCCTTCAGTACAGCCCAACCCACCAACAGCAGTTG 4098

## RESULT 2

US-09-648-281-1

; Sequence 1, Application US/09648281

; Patent No. 6515197

; GENERAL INFORMATION:

; APPLICANT: Pulst, Stefan M.

; TITLE OF INVENTION: Transgenic Animal Model of  
; TITLE OF INVENTION: Neurodegenerative Disease and Methods of Use  
; FILE REFERENCE: P-CE 4336  
; CURRENT APPLICATION NUMBER: US/09/648,281  
; CURRENT FILING DATE: 2000-08-24  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 4481  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (163)...(4099)  
US-09-648-281-1

Alignment Scores:  
Pred. No.: 1,33e-256 Length: 4481  
Score: 6961.00 Matches: 1312  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-648-281-1 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
DB 163 ATGCGCTCAGCGCGCGAGCTCTCGGAGTCCGCGGTGGCCACCGAGTCTCGCGCTTC 222  
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40  
DB 223 GCCGAGCAGGTGGCGCGGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCG 282  
QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60  
DB 283 GCGCGGTGGCG 342  
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaAlaSerAspCysPheGlySerAsn 80  
DB 343 GSCCCCTCTCCCGCGCAGAGCTCGCTCTCCCTCCGCGCTCAGACTGTGTTGGTAGCAAC 402  
QY 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro 100  
DB 403 GCGAAACGCGCGCGCGCGCTTCGCGCGCGCTCCGCGCGCTCTCGGTCTCGCGCGCGCT 462  
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyValaProProAlaAla 120  
DB 463 CCGCGCGCTCTCGTCTGCTCTCTCTCCCTTCGCGCGCGCGCGCGCGCGCGCGCG 522  
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140  
DB 523 CCAACCGCGCTCTCCCGCTCGCGCGCGCGTGGTCCCGCGCGCTTCGCGGTCTCTCTG 582  
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
DB 583 GCGCGCGCGCTTCGCGGTCTCCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCACC 642  
QY 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
DB 643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702  
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762  
QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSerSer 220  
DB 763 AGCGCGCTTCTAGCGTTCGCGCGCGCGCGCGCTTCGCGCGCTTCGCGCTTCCTCG 822  
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240  
DB 823 TCCTCGGCACCGCT 882

QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
DB 883 CTGGCGCAGAGTCAAAACAGTAACAAAGAGCTGCTCAGTCTACGATTTCTTTTATGGA 942  
QY 261 IleTyrAlaAlaMetArgMetValHisLeuLeuThrSerValValGlySerLysCysGlu 280  
DB 943 ATCTATGCAATATGAGGATGGTTTATATCTTACATCAGTTGTTGGCTCCAAATGTGA 1002  
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
DB 1003 GTACAAGTGAATAATGAGGTATATATGAAGAGTGTTTTAAACTTACAGTCCCAAGTGT 1062  
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320  
DB 1063 GATTTGGTACTTGTATGCGCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGT 1122  
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
DB 1123 GAAGAAATAATGAGAGATATTTTGTTCAAATGTTTTCAGACTTTGTTGGTACAGTTTAAA 1182  
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
DB 1183 GATATGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTCTAAA 1242  
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380  
DB 1243 GTGAATGCGAACACAAAGAGAGAGCTCGAGCGCTGGATGCGAGTGAATCAGACCC 1302  
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMet 400  
DB 1303 AATCAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGATCCCAATGATG 1362  
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
DB 1363 TTTTCGATATTAATGAAGAAATTTATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
DB 1423 TATACAGTCCCTTAGAAGAGAGATAACTCAGAAGAAATTTTAAACGGGAAGCAAGGCA 1482  
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
DB 1483 AACCAAGTTCAGAGAGAAATTTGAGTCAAGTCCAGTACAAAGCTCAGTGGCCCTGGA 1542  
QY 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
DB 1543 AATGATGATAGAGTGGAGAGAAATAATACACAGCAGTTACAGAGAAATTCAGTGAACGT 1602  
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500  
DB 1603 GAGGGCAGACATAAACACACTAGGAAAAATAATATATATCTCTCTGGACAAAGAAATAGA 1662  
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
DB 1663 GAAGTCAATATCTCGGGAAGTGGAGAGACAGAAATTCACCGCATATGGCGACCGCTGGATCG 1722  
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCCAATCTCTGGTTCA 1782  
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560  
DB 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTCCCATCTCTCTCTCTCT 1842  
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
DB 1843 CGCCACCTTCTCGCTACCACTAGTCCCAACTCTCTTCCACCTCGGCGACCAACCTCT 1902  
QY 581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600  
DB 1903 ACACGGCGCCCTTCAGGCCCTCTCGGGCCATCCAGACCCCTCTCACCCCTCTCTCTCT 1962

QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
DB 1963 CATGGTCTCCAGCTCGTCTACTATGCGCTAAACGCATGCTCTCAAGAGGGCGCTCCA 2022  
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
DB 2023 AGGATGTCCCAAGGCCAGCGACATCTCGAATACACAGATTTCTGCTGGAGGGT 2082  
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAlaAlaThrPro 660  
DB 2083 TCCATATCCAGTGGCCTAGAAATTTGATCCCAACCCAGCCAGTGAAGCAGCTACTCCT 2142  
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTIPSerSerValValSerGlyValPro 680  
DB 2143 CCAGTAGCAAGGACAGTCCCTCGGGGGAAACGTGTCTCATCGTGTGAGTGGGGTTCCA 2202  
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
DB 2203 AGATTATCCCTAAAACCTCATAGACCAGGTCTCCACAGACAGACAGTATTGGAAATACC 2262  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720  
DB 2263 CCCAGTGGGCCAGTTCTTGCTTCTCCCAAGCTGGTATTATTCCTCAACTGAAGCTGTGGC 2322  
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
DB 2323 ATGCTTATTCAGCTGCATCTCTAGCGCTGTAGTCTCTGATCGAACAGAGCTGTATCC 2382  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGAGCAACTCTCTCTGAGGG 2442  
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
DB 2443 AATAAGAAATATTAACCCATGAACATCACTAGCTTCTCAAAAGCTGAAACAAA 2502  
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
DB 2503 GGTATATCACAGTTGTTCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAG 2562  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
DB 2563 AATGATTTTAGTTACAGGCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAA 2622  
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
DB 2623 AATAGAGGGGAGAAAATCAAGAGATTGATCAAGACAAAATTGAACTGCTAAG 2682  
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
DB 2683 GATTCCTTCAATGAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGC 2742  
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
DB 2743 CCCAGCATTTCCCTTCAATACTTAGTAACACGAGCAGCAAGAGGGACCTGAGGTCACT 2802  
QY 881 SerGlnGlyValGlnThrSerProAlaCysLysGlnLysAspAspLysGluGlu 900  
DB 2803 TCCCAAGGGGTTCAGATTCAGCCCCAGCATGTAAACAGAGAAAGACGATAGAGAGAG 2862  
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
DB 2863 AAGAAAGACGCAGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTC 2922  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
DB 2923 AACCCAGCTTCTCTCTAGCCAAAGCCCTTACTACCCCAACTTCACTCGGCGCTCAA 2982  
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyThrGlnPro 960  
DB 2983 GCACAACTAGCCCATCTATGGTGGGTCAACAGCCAACTCCAGTTTATCTACGCT 3042  
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980

DB 3043 GTTTGTTTGCACAAAATATGATGATTCAGTCCAGTGAGCCAGGGCGTCAACCTTTA 3102  
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
DB 3103 TACCCAAATACCTATGACGCCCATGCGAGTGAATCAAGCAACACATATAGAGCAGTACCA 3162  
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020  
DB 3163 AATATGSCCAACACAGCGGACAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222  
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTySerThrGlnTyrValAlaTy 1040  
DB 3223 GCAGCGGGCCCAACGATTCAGCCACCCACAGCTTACTCCACGCAATATGTTGCCTAC 3282  
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
DB 3283 AGTCTCTCAGCAGTTCACCAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 3342  
QY 1061 HisProHisValTySerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
DB 3343 CATCTCTCATGTCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGCGACCAACA 3402  
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
DB 3403 CAGGCCCGCCCTGGTTAGTATCTTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462  
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
DB 3463 CATGCGATGTATGTCATGTGCCAAATTTACCATACAACAGAGACAAAGCCTTCTTTCTAC 3522  
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
DB 3523 TTTTGCATTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCTAACGCTACCTGTCAC 3582  
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160  
DB 3583 CCACATATCTCCACACCTCAGCTTTCAGTACCTCCCTCAGTGGACAGCAGCAAGCAACAT 3642  
QY 1161 GlyIlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180  
DB 3643 GGTGGAGTCACTCTGACCCAGTCTCTGTTACGACCATCAGCACCGCCCGCCAGGCT 3702  
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
DB 3703 CTCCATCTGGCAGTCCACAGCAGCAGTCCAGCATTTTACCGCGGGCTTGGCGCAACT 3762  
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
DB 3763 CCACCTCTCATGACACTCGCTCCAAACGCGAGTGGCCAGCAATAGTTTCCCGCAGCA 3822  
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
DB 3823 CAACAGACTGTCTTTACGATTCATCTCTCAGTTCAGCGCGGTATACCAACCCACCC 3882  
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
DB 3883 CACATGGCCCGTACCTCAGGCTCATGTACAGTACAGGAATGGTCTCTCTCTCATCAACT 3942  
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
DB 3943 GCCCATGCGCAATGATGCTAATGACGACAGCAGCCCGCGGTCCCCAGCGCGCCCTC 4002  
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
DB 4003 GCTCAAGGTGACCTACAGCCCATTTCCAGTCTCGACAAACAGCGCATTTCCCTATATGACG 4062  
QY 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312  
DB 4063 CACCTCTCAGTACAAGCCCAACCAACAGCAGCTTG 4098

RESULT 3

US-09-707-919A-20

```
; Sequence 20, Application US/09707919A
; Patent No. 6623927
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: Method for detection of human spinocerebellar ataxia 2
; TITLE OF INVENTION: gene variants
; FILE REFERENCE: US 443
; CURRENT APPLICATION NUMBER: US/09/707,919A
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-707-919A-20

Alignment Scores:
Pred. No.: 1,33e-256 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-707-919A-20 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB 163 ATGGCTCAGCGCGCGAGCTCTCGAGTCCGCGGTGGCCAGCTCTCGCGCTTC 222
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
DB 223 GCCGAGCAGCGCGCGCGTTCGCGCGCGCTCGCTCCAGCGCGCGCGCGCGCGG 282
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB 283 GCGCGCGTGGCGCGCGCGCGCGGACGATATCCCTCCGCGCGCGCGCGCGCGCG 342
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB 343 GSCCCCCCTCCCTCCGCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyGlyPro 100
DB 403 GGCACAGCGCGCGCGCGCTTCGCGCGCGCTCCGCGCGCTCTGGTCTCGCGCGGCT 462
QY 101 ProArgProPheValValLeuLeuPheProLeuAlaSerProGlyAlaProProAlaAla 120
DB 463 CCGCGCGCTTCGTCGTCTCTTCCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 523 CCAACCGCGCTCCCGCTCGCGCGCGCGTGGTCCCGCGCGGTTCGCGGTCTCTCTTG 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 583 GCGCGCGCGCTCCCGGTGTCCCGCGCGCGCGTCCGAGCGGTGTATGGGCGCTCACC 642
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 643 ATGTCGTGAAGCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerValSerSer 220
DB 763 AGCGCGCTTCTAGCGTCCCGCGCGCGCGCGCTTCGCGCTTCGCTCTCTCTCTCG 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
DB 1903 ACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1962
```

```
Db 823 TCCTCGGCGCACGGCTCCCTCTCGTGTGTCGCGCGACCTCGGCGGCGGAGGCCCGC 882
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 893 CTGGCAGAGGTGCAACAGAGTAAACAAGAGTAAACAAGAGTAAACAAGAGTAAACA 942
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 943 ATCTATGCAAAATATGAGATGCTTATGAGATGCTTATGAGATGCTTATGAGATGCT 1002
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAAGTGAAGAAATGAGAGTATATGAGAGTATATGAGAGTATATGAGAGTAT 1062
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
DB 1063 GATTGGTACTTGGATGCGGCACATGAGAAAGATGAGAAAGATGAGAAAGATGAGAA 1122
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 1123 GAAGAAATAATGAGAGTATTTTGTTCAAATGTTTTCAGACTTGTGTGTCACAGTTTAAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAA 1242
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGGCGAACAACAAGAGAGAGAGCTGGAGCGCTGGAGTGAAGTGAAGTGAAG 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
DB 1303 AATGAGAACTTGAGGCTTTTGGAAATGACGTAATGATGAGTGGATGCCAATGATATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1363 TTTTCGATATAATGAGAAATATGAGTGTAGTGTGCTACGATATGATGAGTGTATCTTC 1422
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTGCCTTAGAAGAGATAACTCAGAGAAATTTTAAACGGGAGCAAGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACCAAGTATGAGAGAAATTTGAGTCAAGTGCAGTCAAGTGCAGTGCAGTGCAGT 1542
QY 461 AsnAspAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATAGGAGTGAAGAGAAATATACACAGCAGTTCAGAGAAATTTCCAGTGAAC 1602
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500
DB 1603 GAGGGGACAGCATAAACATAGGAGAAATAAATATATCTCTCTGAGCAAGAAATAGA 1662
QY 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAAGTCATATCTGGGAAAGTGGGAGACAGAAATTTCCCGGTATGGGCGACCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGCTCCATGCGCATCAAGATCCACTTCTCACATTTTCAACCCGAAATTTCTGGTTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
DB 1783 GACCAAGAGATGATTAATGAGAGTGTTCCTCGGCGATCGCTTCCCATCTCTCTCTCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
DB 1843 CGCCCACTTCTCGCTACAGTCCAGTCCCAACTCTCTTCCACCTCGGCGACCCACCTCT 1902
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
DB 1903 ACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1962
```



QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
DB	1963	CATGGTTCTCCAGTCTCTCTACTATATGCTTAAACGCATGCTTTCAAGGGGCTCCA	2022
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
DB	2023	AGGATGTCCTCCAAAGGCCAGGACATCTCTGAAATCAAGAGTTTCTCTGGAGGGGT	2082
QY	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAlaAlaThrPro	660
DB	2083	TCCATATCCAGTGGCCCTAGAAATTTGATCCCAACCCAGTCAGTCAGCAGTCTCTCT	2142
QY	661	ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro	680
DB	2143	CCAGTAGCAAGACCAAGTCCCTCGGGGGAAACGTGTCATCAGTGTGCTGAGTGGGGTTCCA	2202
QY	681	ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr	700
DB	2203	AGATTATCCCTTAAACTCATAGCCAGGTCTCCAGACAGAACAGTAGTATTGGAATATCC	2262
QY	701	ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla	720
DB	2263	CCAGTGGGCCAGTCTCTGCTTCTCCCAAGCTGGTATTATTCCACTGAAGCTGTGGC	2322
QY	721	MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr	740
DB	2323	ATGCCTATTCCAGTGCATCTCTAGCGCTGTAGTCTCTGCATCGAACAGAGCTGTATAC	2382
QY	741	ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly	760
DB	2383	CCTTCTAGTGAGGCTAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGG	2442
QY	761	AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys	780
DB	2443	AATAAGAAATATTAACCCATGAACATCACTAGCTTCTCAAAAGCTGAAACAAA	2502
QY	781	GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys	800
DB	2503	GGTATATCACCACTGTTCTGAACATAGAAAAACAGATTGATTTAAAGAAATTTAAG	2562
QY	801	AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys	820
DB	2563	AATGATTTTAGTTACAGCAGCAAGTCTTACTTCTGAATCTATGGATCAACTATAAACAA	2622
QY	821	AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys	840
DB	2623	AATAGAGAGGGAGAAATCAAGAGATTGATCAAGACAAAATTGAAACCAAGTGTCTAAG	2682
QY	841	AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer	860
DB	2683	GATTCITTCATTGAAATATAGCAGCAGCAACTGTACCACTGGCAGCAGCAAGCCGAATAGC	2742
QY	861	ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr	880
DB	2743	CCCAGCATTTCCCTCTCAATACTTAGTAACACCGAGCACAAGAGGGACCTGAGGTCACT	2802
QY	881	SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu	900
DB	2803	TCCCAAGGGGTTAGACTTCCAGCCCGCATGTAAACAGAGAAAGCATAGAGAGAG	2862
QY	901	LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe	920
DB	2863	AAGAAAGACGCGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTTC	2922
QY	921	AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln	940
DB	2923	AACCCAGGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGCGCTCAA	2982
QY	941	AlaGlnProSerProSerMetValGlyHisGlnProThrProValThrThrGlnPro	960
DB	2983	GCACAACTAGCCCATCTATGGTGGGTGTCATCAACAGCCAACTCCAGTTTATCTCAGCCT	3042

RESULT 4

QY	961	ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980
DB	3043	GTTTGTGTCACCAAAATATGATGATCAGTCCAGTCAGCCAGCGCGTGCACACCTTTTA	3102
QY	981	TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro	1000
DB	3103	TAACCAATACCTATCAGCCCATGCGAGTGAATCAAGCAACACATATAGAGCAGTACCA	3162
QY	1001	AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer	1020
DB	3163	AAATATGCCCAACAGCGGCAAGACCATCATCAGAGTGCCATGATGACCAACGCGTCA	3222
QY	1021	AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr	1040
DB	3223	GCAGCGGGCCCAACCATTTGCAGCCACCCACAGCTTACTCCAGCAATATGTTGGCTTAC	3282
QY	1041	SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln	1060
DB	3283	AGTCTCTAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAG	3342
QY	1061	HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProThr	1080
DB	3343	CATCTCTATGTCATAGTCTCTTAATACAGGGTAATGCTAGAAATGATGCGCACCAACA	3402
QY	1081	HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr	1100
DB	3403	CACGCCCAACCTGGTGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACG	3462
QY	1101	HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr	1120
DB	3463	CATGCGATGATGTCATGTCCTCCAAATTTACCATACACAGAGACAAAGCCCTTCTTTCTAC	3522
QY	1121	PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis	1140
DB	3523	TTTGGCATTTCCACGGGCTCCCTTGCTCAGCAGTATGCGACCTAACGCTACCTCGCAC	3582
QY	1141	ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis	1160
DB	3583	CCACATCTCCACACCTCAGCCTTTCAGTACCTCCACCTGGACAGCAGCAAGCCCAACAT	3642
QY	1161	GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla	1180
DB	3643	GGTGGAGTCATCTGTCACCCAGTCTCTGTCAGCACCATCAGCAGCCGCGCCAGGCT	3702
QY	1181	LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr	1200
DB	3703	CTCCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTACACGCGGGCTTGGCCCAACT	3762
QY	1201	ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla	1220
DB	3763	CCACCTCTCATGACACTGCTCTCCAAACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCA	3822
QY	1221	GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro	1240
DB	3823	CAACAGACTGTCTTTACGATCCATCTCTCAGCTTCAGCCGCGTATACCAACCCACCC	3882
QY	1241	HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr	1260
DB	3883	CACATGGCCCACTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCTCAACT	3942
QY	1261	AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyProGlnAlaAlaLeu	1280
DB	3943	GCCCATGCGCAATGATGCTAATGACGACACAGCCACCCGCGGCTCCCAAGCGCGCTC	4002
QY	1281	AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr	1300
DB	4003	GCTCAAGTGCACTACAGCCCATTCAGTCTCGAACAAACAGCGCATTTCCCTATATGACG	4062
QY	1301	HisProSerValGlnAlaHisHisGlnGlnGlnLeu	1312
DB	4063	CACCTTTCAGTACAAAGCCCAACCAACACAGCAGTTG	4098

US-09-083-268-2  
 ; Sequence 2, Application US/09083268  
 ; Patent No. 6673535  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pulst, Stefan M  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR  
 ; TITLE OF INVENTION: ATAXIA-2 AND PRODUCTS RELATED THERE TO  
 ; NUMBER OF SEQUENCES: 18  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Mueeting, Raasch & Gebhardt, P.A.  
 ; STREET: 119 No. 6673535th Fourth Street  
 ; CITY: Minneapolis  
 ; STATE: Minnesota  
 ; COUNTRY: USA  
 ; ZIP: 55401  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/083,268  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/727,084  
 ; FILING DATE: 08-OCT-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McCormack, Myra H  
 ; REGISTRATION NUMBER: 36,602  
 ; REFERENCE/DOCKET NUMBER: 232.00010101  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 612/305-1220  
 ; TELEFAX: 612/305-1228  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4481 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: both  
 ; TOPOLOGY: both  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 163...4101  
 ; US-09-083-268-2

Alignment Scores:  
 Pred. No.: 1,33e-256 Length: 4481  
 Score: 6961.00 Matches: 1312  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-802-228-2 (1-1312) x US-09-083-268-2 (1-4481)  
 QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
 DB 163 ATCGCTCAGCGCGCAGCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTC 222  
 QY 21 AlaAlaAlaArgTTPProGlyTTPArgSerLeuGlnArgProAlaArgSerClyArg 40  
 DB 223 GCCGAGCAGGTGCCCGGGTGGCGCTCGCTCCAGCGCGCGCGCGCGGCGGCGG 282  
 QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProPro 60  
 DB 283 GCGCGGGTGGCGGCGCGCGCGGACGATATCCTCGCGCGCCCTCCCGCGCGCGCGG 342  
 QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80  
 DB 343 GCGCCCGCTCCCTCCCGCGCAGAGCTCGCGCTCCCTCCGCTCAGACTGTTTTGGTAGCAAC 402  
 QY 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyGlyPro 100

Db 403 GGCAACGGCGCGCGCGCTTTCGGCCCGGCTTCCGGCGGCTCTCTGGTCTCGCGGGCGCT 462  
 QY 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120  
 DB 463 CCGCGCGCTTTCGTCGTCTCTTCTCCCTTCGCCAGCCCGGGCGGCTTCCGGCGCGG 522  
 QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProAlaArgSerGlyValSerLeu 140  
 DB 523 CCAACCGCGCGCTCCCGCTCGGCGCGCTCGCTCCCGCGCGCTTCCGGCGCTCTCTTG 582  
 QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
 DB 583 GCGCGCGCGCTCCCGGCTTCCCGCGCGCGCGCTGCGAGCGGTGTATGGGCGCGCTCAC 642  
 QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
 DB 643 ATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702  
 QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
 DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762  
 QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerProSerSerSerValSerSer 220  
 DB 763 AGCGCGCTTCTAGCGTCCCGCGCGCGCGCTTTCGGCGCTCTCGCTCTCGCTCTCG 822  
 QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240  
 DB 823 TCCTCGGCGCAGGCTCCCTCTCGTGGTCCGCGCGACCTCCGGCGCGGCGGAGCGCGC 882  
 QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
 DB 883 CTGGCGCAGAGGTGCAAAACAGTAACAAAGGACTGCTCAGTCTAGCATTTCTTTGATGGA 942  
 QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
 DB 943 ATCTATGCAAAATATGAGGATGGTTCATATCTACATCAGTTGTGGCTTCCAAATGTCAA 1002  
 QY 281 ValGlnValLysAsnGlyGlyIleTyrGlyValPheLysThrTyrSerProLysCys 300  
 DB 1003 GTACAAGTGAAATGGAGGTATATATGAAGAGATTTTAAACCTTACAGTCCGAAGTGT 1062  
 QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320  
 DB 1063 GATTTGGTACTTGTATGCGGCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGT 1122  
 QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys 340  
 DB 1123 GAAGAAATAATGGAGAGATTTTGTTCAAATGTTTCAGACTTTGTGGTACAGTTTAA 1182  
 QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
 DB 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAA 1242  
 QY 361 ValAsnGlyGluHisLysGlyLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380  
 DB 1243 GTGAATGGCGAACAACAAGAGAGAGCCTGGAGCGCTGGGATGCGAGGTGAATCAGACCC 1302  
 QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
 DB 1303 AATCAGGAACCTTAGGCTTTTGGAAATGACGTATCTAATGATGGATGCCAATGATATG 1362  
 QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
 DB 1363 TTTTCGATATATGAAGAAATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
 QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
 DB 1423 TATCAGTGCCCTTAGAAAGAGATAACTCAGAAAGATTTTAAACCGGGAAGCAAGGCA 1482  
 QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460

1483 AACAGTTAGCAGAGAAATTGAGTCAAGTCCAGTACAAAGCTCGAGTGCCTGGNA 1542  
461 AsnAspAspArgSerGluGluGluGlySerThrAlaValGlnArgAsnSerSerGluArg 480  
1543 AATGATGATAGGAGTGAGGAGAGAAATACACAGCAGATTTCAGAGAAATCCAGTGAACGT 1602  
481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
1603 GAGGGGCACAGCATAAACACTAGGGAATAAATATATTCCTCGGACAAAGAAATAGA 1662  
501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
1663 GAAGTCATATCTGGGAAGTGGAGACAGAAATTCACCGCGTATGGCCAGCGCTGGATCG 1722  
521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
1723 GGCTCCATGCCATCAAGATCCACTCTCACACTTCAGATTCAACCCGAATTCCTGGTTCA 1782  
541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560  
1783 GACCAAGAGTAGTTAATGAGGTGTTCCTGGCCATCGCCTTGCCCATCTCTTCCTCT 1842  
561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
1843 CGCCCACTTCTCGCTACCACTGAGTCCCACTCTTCCACCTCGGCAGCACCCTCT 1902  
581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
1903 ACAGGCGCCCTCCAGGCGCCCTCGGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
1963 CATGGTTCTCCAGCTCTGCTCTACTATGCTTAAACGCGATGCTTCAGAGGGGCTCCA 2022  
621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
2023 AGGATGTCCCAAGGCCAGCACATCTCTGAAATCACAGATTTCTGCTGGAGGGGT 2082  
641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
2083 TCCATATCCAGTGGCTAGAAATTTGATCCACACCCAGTGAAGCAGTACTCTCT 2142  
661 ProValAlaArgThrSerProSerGlyGlyThrTyrSerSerValValSerGlyValPro 680  
2143 CCAGTAGCAGGACCAAGTCCCTCGGGGAAACGTGTGTCATCAGTGTGCGGTTCGA 2202  
681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
2203 AGATTATCCCTTAAACTCATAGCCAGGTCTCCAGACAGAAACAGTATTTGAAATACC 2262  
701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720  
2263 CCAGTGGGCCAGTTCTTGCTTCTCCCAAGCTGGTATTTCCAACTGAGCTGTGCC 2322  
721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
2323 ATGCTTATTCAGCTGCATCTCTACGCTGTAGTCTCTGCATCGAACAGAGCTGTACC 2382  
741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
2383 CCTTCTAGTAGGCTAAAGATTCCAGGCTTCAGAGTCAGAGGAGAGAACTCTCTCGAGGG 2442  
761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
2443 AATAAGAAATATTAACCCATGAACATCACCTAGCTTCTCAAAAGCTGAACAAA 2502  
781 GlyIleSerProValValSerGluHisArgLysGlnIleAspLeuLysLysPheLys 800  
2503 GGTATATCACAGTTGTTCTGAACATAGAAAAACAGATTGATTTAAAGAAATTAAG 2562  
801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuAsnLys 820  
2563 AATGATTTTAGTTACAGCCCAAGTTCTTCTGAAATCTATGATCAACTACTAAAAAAA 2622

821 AsnArgGluGlyGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
2623 AATAGAGGGAGAGAAATCAAGAGATTGTATCAAGAGCAAAATTTGAACCAAGTCTTAAG 2682  
841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
2683 GATTCTTTTCAATTAAGATAGCAGCAACTGTACCACTGGCAGCAGCAAGCCGAATAGC 2742  
861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
2743 CCCAGCATTTCCCTTCAATTAAGTAAACGAGGAGCAAGAGGGGACCTTGAAGTCACT 2802  
881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900  
2803 TCCCAAGGGGTTACAGTTCAGCCCGCAGCATGTAAACAGAGAAAGACATTAAGGAAGAG 2862  
901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
2863 AAGAAAGACGACGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTTC 2922  
921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
2923 AACCCAGCTGCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCTCAA 2982  
941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
2983 GCACAACCTAGCCCATCTATGTGTGGTCAACAACAGCCCAACTCCAGTTTATATCTCAGCCT 3042  
961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
3043 GTTTGTTTGGACCAAAATATGATGTATCCAGTCCCAAGTCCAGCCGCGGTGCAACCTTTA 3102  
981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
3103 TACCCAATACCTATGAGGCCCATGCGAGTGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162  
1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020  
3163 AATATGCCCAACAGCGGCAAGCAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222  
1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
3223 GAGCGGGGCCCAACCATTCAGCCACCCCAAGTACTTCCAGCAATATGTTGCTTAC 3282  
1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
3283 AGTCTCTACAGCTTCCCAATCAGCCCTTGTTCAGCATGTGCCATATTCAGTCTCAG 3342  
1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
3343 CATCTCATGTCTATAGTCTCTGTAATACAGGGTAAATGTAGATGATGCGACCCCAACA 3402  
1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
3403 CAGCCCAACCTCGTTTGTATCTTTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462  
1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
3463 CATGCGATGTATGCTATGCTCCCAATTTACCATACAAAGAGAGACAAGCCCTTCTTCTAC 3522  
1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
3523 TTTGCCATTTCCAGGGCTCCCTTGTCTCAGCAGTATGCGACCCCTAACGCTACCTCGCAC 3582  
1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160  
3583 CCACATATCCACACCTCAGCCTTCCAGTACCCCACTGGACAGCAGCAAGCCCAACAT 3642  
1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180  
3643 GGTGGAAGTCACTCTCGACCCAGTCTGTTTTCAGCACCATCAGCACCCGCGCCAGGCT 3702

Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTACGATTTTACCACTGGCGGCTTGGCGCAACT 3762  
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
Db 3763 CCACCTCATGACACCTTGGCTTCCAAACAGCAGTCCGACAGAAATGATTTTCCACGACGA 3822  
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
Db 3823 CAACAGACTGTCTTTACGATCCATCTTCTCAGGTTTCCAGCGGCTATACCAACCCACCC 3882  
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
Db 3883 CACATGGCCCACTGATCTTCCAGCTCATGATGAGTCCAGAAATGTTTCTTCTCATCCAACT 3942  
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
Db 3943 GCCCATGGCCCATGATCTATATGACACACAGCCAGCGGCTCCAGCGCCCTC 4002  
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
Db 4003 GCTCAAGTGCACTACAGCCATCTCCAGTCTCGACACAGCGCATTTTCCCTATATGAGC 4062  
Qy 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312  
Db 4063 CACCCCTTCAGTACAAGCCACCCACCAACAGCAGTTG 4098

RESULT 5  
US-08-981-998A-2  
; Sequence 2, Application US/08981998A  
; Patent No. 6844431  
; GENERAL INFORMATION:  
; APPLICANT: PULST, STEFAN M.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCERECELLAR  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.  
; STREET: 119 NORTH FOURTH STREET, SUITE 203  
; CITY: MINNEAPOLIS  
; STATE: MINNESOTA  
; COUNTRY: 55401  
; ZIP: 55401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 11-May-1998  
; APPLICATION NUMBER: US/08/981,998A  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 97/42314  
; FILING DATE: 08-MAY-1997  
; APPLICATION NUMBER: US 08/727,084  
; FILING DATE: 08-OCT-1996  
; APPLICATION NUMBER: US 60/022,207  
; FILING DATE: 19-JUL-1996  
; APPLICATION NUMBER: US 60/017,388  
; FILING DATE: 08-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MUETING, ANN M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 232.00010120  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-305-1217  
; TELEFAX: 612-305-1228  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4481 base pairs  
; TYPE: nucleic acid

; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 163..4101  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-08-981-998A-2  
Alignment Scores:  
Pred. No.: 1,33e-256 Length: 4481  
Score: 6961.00 Matches: 1312  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
US-10-802-228-2 (1-1312) x US-08-981-998A-2 (1-4481)  
Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
Db 163 ATGGCTCAGCGCGCGCAGCTCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTC 222  
Qy 21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg 40  
Db 223 GCCGACGACAGTGGCGCGGCTGGCTCGCTCCAGCGCGCGCGCGCGCGCGGCGG 282  
Qy 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60  
Db 283 GCGCGCGGTGGCGCGCGCGCGGACCGTATCCCTCCCGCGCGCTCCCGCGCGCGCG 342  
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80  
Db 343 GGCCCCCTCCCTCCCGCGCAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAAC 402  
Qy 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100  
Db 403 GGCACACGCGCGCGCGGCTTTTCGCGCGCGGCTCCCGCGCGGCTCTCGCGCGGCGCT 462  
Qy 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120  
Db 463 CCGCGCGCTTCGTGCTGCTCTTCTCCCTCCCGCGCGCGCGCGCGCGCGCGCGCG 522  
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140  
Db 523 CCAACCGCGCGCTCCCGCTCGCGCGCGTGGCTCCCGCGCGCTTCGCGCGCTCTCTTG 582  
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
Db 583 GCGCGCGCGCTCCCGCTGTGTCCCGCGCGCGGTGCGAGCGGCTGTATGGCGCGCTCACC 642  
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
Db 643 ATGCTGCTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702  
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
Db 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762  
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerProSerSerValSerSer 220  
Db 763 AGCGCGCTTCTAGGCTGCGCGCGCGCGCGCGCTTTCGCGCGCTCTCGCTCTCTCTCG 822  
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyValArgProGly 240  
Db 823 TCCTCGCGCGCGCTCCCTCTCTCGGTGCTCGCGCGACCTCCCGCGCGCGGAGCGCGCG 882  
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
Db 883 CTGGGACAGGTCGAAACAGTAAACAAAGGACTGCGCTCAGTCTACGATTTCTTTTATGGA 942  
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280

Db 943 ATCTATGCAAAATATGAGGATGGTTCATATACATCACTGTTGGCTCCAAATGTGAA 1002  
Qy 281 ValGlnValIleAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
Db 1003 GTACAGTGAATAATGAGGTATATATGAAGAGATTTTAAACTTACAGTCCGAGTGT 1062  
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320  
Db 1063 GATTTGGTACTTGTATCCGACATGAGAAAGTACAGAAATCCAGTTCCGGGCGCAACGT 1122  
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
Db 1123 GAAGAATAATGAGAGATTTTGTTCAAATGTTACAGCTTTGTGTGATCAGTTTAA 1182  
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
Db 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTCTATCAGTGTCTAAA 1242  
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380  
Db 1243 GTGAATGGCGAAACACAAAGAGAGAGACCTGGAGCCCTGGGATGCAAGTGAACCTCACAGCC 1302  
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
Db 1303 AATGAGGAACCTTGAGGCTTTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATG 1362  
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
Db 1363 TTTGATATATGAGAAATTTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
Db 1423 TATACAGTGCCTTAGAAAGAGATAACTCAGAAGATTTTAAACCGGAAGCAAGGGCA 1482  
Qy 441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrIleValAlaLeuGlu 460  
Db 1483 AACCAAGTACGACGAAGAAATTTGATCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAA 1542  
Qy 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
Db 1543 AATGATGATAGGATGGGAAGAAATAACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602  
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500  
Db 1603 GAGGGGCACAGCATAACACTAGGAAATAATAATATATCTCTCTGGACAAAGAAATAGA 1662  
Qy 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
Db 1663 GAAGTCAATATCCTGGGAAGTGGGAGACAGAAATTCACCGGTATGGGCCAGCCTGGATCG 1722  
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
Db 1723 GGCTCCATGTCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGGTTCA 1782  
Qy 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560  
Db 1783 GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGGCCATCTCTCTCTCTCT 1842  
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
Db 1843 CGCCCACTTCTGCTACCAAGTCAAGTCCCAACTCTCTCCACCTCGGGCAGCCACCCCT 1902  
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
Db 1903 ACAGGGCCCTCCAGCCCTCCGGCCATCCAGACCCCGTCTCACCCCTCTGCT 1962  
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Db 1963 CATGTTCTCAGTCTCTGCTCTACTATGCTTAACGCATGCTTTCAGAAAGGSCCTCCA 2022  
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Db 2023 AGGATGTCCCAAGGCGCCAGCGACATCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT 2082

Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
Db 2083 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACACCCAGTGAACAGCTACTCTCT 2142  
Qy 661 ProValAlaAaThrSerProSerGlyGlyThrTyrTrpSerSerValValSerGlyValPro 680  
Db 2143 CCAGTAGCAAGACACAGTCCCTCGGGGAACTGGTCACTCAGTGGTCACTGGGTTCCTCA 2202  
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
Db 2203 AGATTATCCCTTAAACTCATAGACCCAGGTCTCCAGACAGAACAGTATTGCAATACC 2262  
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
Db 2263 CCCAGTGGGCCAGTTCTTCTCTCCCAAGCTGGTATTATTCCAACTCAAGCTGTTCCT 2322  
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
Db 2323 ATGCCTATTCCAGCTGCATCTCTACGCTGTAGTCTCTGCATCGAACAGAGCTGTACC 2382  
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
Db 2383 CCTTCTAGTAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGGG 2442  
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
Db 2443 AATAAAGAAAAATTTAAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAA 2502  
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800  
Db 2503 GGTATATCACCAAGTTGTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2562  
Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
Db 2563 AATGATTTTAGGTACAGCCAGTTCTACTTCTGAATCTATGATCACTACTATAACAAA 2622  
Qy 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
Db 2623 AATAGAGAGGAGAAAAATCAAGAGATTTCATCAAAAGACAAAAATGAAACCAAGTCTAAG 2682  
Qy 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
Db 2683 GATTCCTTTTCAATGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCAATAGC 2742  
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
Db 2743 CCAGCATTTCCCTTCAATACTTATTAGTACACGAGCACAAGAGGGGACCTGAGGTCACT 2802  
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900  
Db 2803 TCCCAAGGGGTTCCAGACTTCCAGCCCGAGCATGTAAACAAGAGAAAGACGATAAGAAAGAG 2862  
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
Db 2863 AAGAAAGACGACGCTGAGCAAGTTAGAAATCAACATTGAATCCCAATGCAAGAGAGTTC 2922  
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
Db 2923 AACCCAGTTCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCCCTCGGCCTCAA 2982  
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
Db 2983 GCACAACTTAGCCCACTATGTTGGGTTCATCAACAGCCAACTCCAGTTTATCTACGCTC 3042  
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
Db 3043 GTTTGTTTTCACCAAAATATGATGTATCCAGTCCCAAGTGGCCAGCGCTGCAACCTTTA 3102  
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
Db 3103 TACCAATATCTATGAGCCCATGCTCCAGTGAATCAAGCAAGACATATAGAGCAGTACCA 3162

```
QY 1001 AenMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCCAACAGCGGACAGCAGCATCATCAGAGTGCCATGATCACCACGCGTCA 3222
QY 1021 AlaAlaGlyProPheAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGGGGGCCACCATGTCAGCCACCCACAGAGTTACTCCACCAATATGTTGCCATC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCTCTCATGTCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGGCACCAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyValHisGlnGlnThr 1100
Db 3403 CACGCCCGCCGCTGGTTAGTATCTTCTTCAGCACTCAGTACGGGGCTCATGAGCAGAGC 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGCGATGTATGTCATGCTCCCAATACCATACACAGGAGACAAAGCCCTTCTTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGCGCATTTCCACGGGCTCCCTTCTTCAGCAGTATGGCGACCCCTAACGCTACCCCTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 3583 CCACATATCTCCACACCTCAGCTCCTCAGTACCCCTCCTGGACACAGCAAGCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Db 3643 GGTGGAAGTCATCTCTCCACCCAGTCTCTGTTCCAGCACCATCAGCACCGCGCCCGCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCCATTTACCACGCGGGGCTTGCGCCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3763 CCACCCCTCATGACACCTGCTCCAAACAGCAGTCCGACAGAAATAGTTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3823 CAACGAGCTGTCTTTACGATCCATCTCTCAGCTTCAGCGGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGGGCCACGTCACCTCAGGCTCATGTACAGTCAGGAATGGTTCTCTCATCCAACT 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGGCCCAATGATGCTAATGACGACACAGCCACCCGGGGTCCCCAGCGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrAlaHisPheProTyrMetThr 1300
Db 4003 GCTCAAGTGCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4063 CACCCCTTCAGTACAGGCCCAACCAACACAGCAGTTG 4098
```

RESULT 6

```
US-09-949-016-4498
; Sequence 4498, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4498
; LENGTH: 4484
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4498

Alignment Scores:
Pred. No.: 5,87e-254 Length: 4484
Score: 6891.50 Matches: 1309
Percent Similarity: 99.70% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.00% Indels: 3
DB: 3 Gaps: 1
```

US-10-802-228-2 (1-1312) x US-09-949-016-4498 (1-4484)

```
QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGGGCTCAGCGGGCGCAGCTCTCCGAGTCCCGGGTGGCCACGAGTCTCGCG- TTC 221
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db 222 GCCGACGACAGTGGCGGGTGGCTCGTCCAGCGCGCGCGCGCGGAGCGGCGG 281
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 282 GCGCGCGTGGCGGGCGCGGACCGTATCCCTCCGCGCGCTTCTCCCGCGCGCCC 341
QY 61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 342 GCGCGCGCTCCCTCCCGCAGAGCTCCCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 401
QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
Db 402 GGCAACGGCGCGCGGCTTTCGGCCCGGCTCCCGCGGCTCTTGTTGTCGCGCGGCT 461
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
Db 462 CCGCGCGCTTCTGTCGCTCTCTTCTCCCTCCGACCGCGCGCGCGCTCCCGCGCGG 521
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 522 CCAACCGCGGCTCCCGCTCGGCGCGCGCGCTCCCGCGCGCTTCGGGCTCTCTCTTG 581
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 582 GCGCGCGGCTCCCGGCTGTCGCCCGCGCGCTGCGAGCGGCTGATGGCCCCCTCAC 641
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 642 ATGTGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 701
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 702 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 761
QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerValSerSer 220
Db 762 AGCGGCTTCTAGCGTCCCGCGCGCGCGCTTCCGCGCTCTCGCTCTCTCTCTCTCG 821
QY 221 SerSerAlaThrAlaProSerSerValAlaAlaThrSerGlyGlyGlyArgProGly 240
Db 822 TCCTCGGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 881
```

241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
882 CTGGGCAGAGGTCGAAACACAGTAAACAGACTGCCTCAGTCTACGATTTCTTTTATGGA 941  
261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
942 ATCTATGCAAAATATGAGGATGGTTCAATATCTTACATCATGTTGTGGCTCCAAATGTGAA 1001  
281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
1002 GTACAAGTGAAATATGAGGTATATATGAAGAGTTTAAACCTTACAGTCCGAAGTGT 1061  
301 AspLeuValLysAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320  
1062 GATTTGGTACTTGTATGTCGCCACATGAGAAAGTACAGAAATCCAGTTCGGGGCGGAAACGT 1121  
321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
1122 GAAGAAATATGAGAGTATTTTGTCAAAATGTTTCAAGCTTTGTGTGATCAGTTTAA 1181  
341 AspMetAspSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
1182 GATATGGACTCCAGTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTTAA 1241  
361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380  
1242 GTGAATGGCGCAACACAAAGAGAGGACCTGGAGCCCTGGGATCGAGGTGAATCTCAGCC 1301  
381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
1302 AATGAGAACTTGGAGCTTTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATG 1361  
401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
1362 TTTTCGATATATGAAGAAATATGTTGTGTAGTGTCTACGTATGATGAGTATTAICTTG 1421  
421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
1422 TATACAGTGCCTTAGAAAGAGATACTCAGAAGAAATTTTAAACCGGAAGCAAGGGCA 1481  
441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
1482 AACCACTTAGCAGAAATATGATCAAGTGCCAGTACAAAGCTCGAGTGCCCTGGAA 1541  
461 AsnAspAspArgSerGluGluGlyLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
1542 AATGATGATAGGATGAGGAGAGAAATATACAGAGCTTCAGAGAAATTCAGTGNACGT 1601  
481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500  
1602 GAGGGGCACAGCATAAACCTAGGGAAATATAATATATCTCTCTCGACAAAGAAATAGA 1661  
501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
1662 GAAGTCATATCTCTGGGGAATGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCG 1721  
521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
1722 GGTCTCATGTCATCAAGATCCATCTTCACTTCAGATTTCAACCCGGAATTCGTGTCA 1781  
541 AspGlnArgValValAsnGlyGly----ValProTyrProSerProCysProSerProSe 559  
1782 GACCAAGAGTAGTTAATGAGGCAAGTGTTCCTCGGCCATCGCTTGGCCATCTCCTTC 1841  
559 rSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaIle 579  
1842 CTCTCGGCCACCTTCTGCTACAGTCAGGTCCCAACTCTCTTCCACCTCGGGCAGCCAC 1901  
579 rProThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSe 599  
1902 CCCTACAGCGCCGCTTCCAGGCCCTTCCAGGCCATCCAGACCCCTCCGCTCACCCTC 1961

599 rAlaHisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyPr 619  
1962 TGCTCATGGTCTCAGCTCTCTCTACTATGCTTAACGCATGTCTTTCAGAGGCC 2021  
619 oProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyAr 639  
2022 TCCAAAGATGTCCCAAGGCCAGCGACATCTCGAAATCACAGAGTTCTGTCTGGCAG 2081  
639 gGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaTh 659  
2082 GGGTTCCATATCCAGTGGCTAGAAATTTGTATCCACAAACCCAGCTGAAGCAGCTAC 2141  
659 rProProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyVa 679  
2142 TCCTCCAGTAGCAAGGACAGTCCCTCGGGGGAAACGTGGTTCATCAGTGGTTCAGTGGGT 2201  
679 lProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAs 699  
2202 TCCAAAGATTTATCCCTTAAACTCATAGACCCAGGTCTCCACAGACAGACAGTATTGGAAA 2261  
699 nThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaVa 719  
2262 TACCCCCAGTGGGCAGTCTTGTCTCTCCCAAGCTGGTATTATTCCAACCTGAAGCTGT 2321  
719 lAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnAtGAlaVa 739  
2322 TGCCATGCTTATTCAGCTGCATCTCTACGCTCTAGTCTCTGCATCGAACAGAGCTGT 2381  
739 lThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAl 759  
2382 TACCCCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGCGAGAACTCTCTGC 2441  
759 aGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAs 779  
2442 AGGGAATAAGAAAAATATTAAACCCCAATGAAACATCAGCTAGCTTCTCAAAAGCTGAAA 2501  
779 nLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPh 799  
2502 CAAAGGTATATCACCAGTTGTTTCTGAACATAGAAAAACAGATTGATTTTAAAGAAAT 2561  
799 eLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAs 819  
2562 TAAGAAATGATTTAGTTAGTACAGCCAAAGTTCTACTTCTGAATCTATGGATCAACTATA 2621  
819 nLysAsnArgGluGlyGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAl 839  
2622 CAAAAATAGAGAGGAGAAAAATCAAGAGATTGATCAAAAGCAAAATTTGAACCAAGTGC 2681  
839 aLysAspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAs 859  
2682 TAAGGATTTCTTTTCAATTTGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAA 2741  
859 nSerProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluVa 879  
2742 TAGCCCCAGCATTTCCCTTCAATATCTAGTAAACAGGAGCAGCAAGAGGGGACCTGAGGT 2801  
879 lThrSerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGly 899  
2802 CACTTCCCAAGGGTTTCAGACTTCCAGCTCCAGCCAGCATGTAAACAAAGAGAAAGCATAAGGA 2861  
899 uGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGly 919  
2862 AGAGAAAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTTGAAATCCCAATGCAAGGA 2921  
919 uPheAsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgPr 939  
2922 GTTCAACCCAGCTTCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCCACCTCGGCC 2981  
939 oGlnAlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGly 959  
2982 TCAAGCACAACTAGCCCATCTATGTGGGTCTATCAACAGCCAACTCCAGTTTATACTCA 3041  
959 nProValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnPr 979



```
Db 3042 GCCTGTTGTTGGACCAAAATATGATGATCCAGTCAGCCAGGCGTGCACCC 3101
QY 979 oLeuTyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaVa 999
Db 3102 TTATACCAATACCTATGACGCCCATGCGAGTGAATCAAGCAAGACATATAGAGCAGT 3161
QY 999 lProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAl 1019
Db 3162 ACCAAATATGCCCAACAGCGCAAGACAGCATCATCAGAGTGCATGATGCACCCAGC 3221
QY 1019 aSerAlaLaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAl 1039
Db 3222 GTCAGAGCGGGCCCAACCAATGTCAGCCACCCACAGCTTACTCCAGCAATATGTTGC 3281
QY 1039 aTyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSe 1059
Db 3282 CTACAGTCTCAGCAGTTCGCCAAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTC 3341
QY 1059 rGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProPr 1079
Db 3342 TCAGCATCTCATGTCTATAGTCCCTGTAATACAGGGTAAATGCTAGAAATGATGGCACCC 3401
QY 1079 oThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGl 1099
Db 3402 AACACAGCGCCAGCCTGGTTAGTATCTTTCAGCACTCAGTACGGGCTCATGAGCA 3461
QY 1099 nThrHisAlaMetTyrAlaCysProGlyLeuProTyrAsnLysGluThrSerProSerPh 1119
Db 3462 GACGCATGCGATGTATGATGTCCCAATTTACCATACAAACAGGAGACAGCCCTCTTT 3521
QY 1119 eTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrIle 1139
Db 3522 CTACTTTGCCATTCTCAGCGGCTCCTTGTCTCAGCAGTATGCGCACCCCTTAAGCTTACCCT 3581
QY 1139 whiSerProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGl 1159
Db 3582 GCACCCACATATCTCACACCTTCAGCTTCAGTACCCCTCATGACAGCAGCAAGGCA 3641
QY 1159 nHisGlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGl 1179
Db 3642 ACATGTTGGAAGTCATCTCGACCCAGTCTCTTCAGCACCATCAGCACCAGCGCGCCA 3701
QY 1179 nAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaPr 1199
Db 3702 GGCTCTCCATCTGGCCAGTCCACAGCAGCAGTTCAGCCATTTACCACGCGGGCTTGCGCC 3761
QY 1199 oThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAl 1219
Db 3762 AACTCCACCTCCATGACACCTGCTCCAAACGCGAGTCGCCACAGATAGTTCCTCCAGC 3821
QY 1219 aAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPr 1239
Db 3822 AGCACAAACAGACTGCTTTACGATCAATCTTCTCAGCTTCAGCGCGGTATACCAACC 3881
QY 1239 oProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisPr 1259
Db 3882 ACCCCACATGGGCCACGTACTCCTCAGGCTCATGTACAGTCAGGAATGTTCTCTTCATCC 3941
QY 1259 oThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAl 1279
Db 3942 AACTGCCATGGCCAAATGATGCTAATGACGACACAGCCACCCGCGGTGCCAGCGCGC 4001
QY 1279 aLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMe 1299
Db 4002 CCTCGCTCAAAGTGCACATACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTATAT 4061
QY 1299 tThrHisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4062 GACGACACCTCTCAGTACAGGCCACCCACCAACAGCAGATTG 4101
```

RESULT 7

US-08-981-998A-4

```
; Sequence 4, Application US/08981998A
; Patent No. 6844431
; GENERAL INFORMATION:
; APPLICANT: PULST, STEFAN M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEITING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: 55401
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,998A
; FILING DATE: 11-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 97/42314
; FILING DATE: 08-MAY-1997
; APPLICATION NUMBER: US 08/727,084
; FILING DATE: 08-OCT-1996
; APPLICATION NUMBER: US 60/022,207
; FILING DATE: 19-JUL-1996
; APPLICATION NUMBER: US 60/017,388
; FILING DATE: 08-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..3457
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-981-998A-4
Alignment Scores:
Pred. No.: 4,02e-205 Length: 3798
Score: 5606.50 Matches: 1063
Percent Similarity: 94.20% Conservative: 41
Best Local Similarity: 90.70% Mismatches: 37
Query Match: 80.54% Indels: 31
DB: 3 Gaps: 8
US-10-802-228-2 (1-1312) x US-08-981-998A-4 (1-3798)
QY 148 ProArgProAlaAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysProGlnGln 167
Db 11 CCCCAGCGCGCGTGGAGCGGTGTATGGCCGCTCACCATGTCGCTGAACCCGACGCG 70
QY 168 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 187
Db 71 CAG----- 73
QY 188 ProProProAlaAlaAlaAsnValArgLysProGlyGlySerGlyLeuAlaSerPro 207
Db 74 ---CCGCGCGCGCGCGCCACTGGCGCGCAAGCCCGCGGC---GGCCTGCTCTGTCGCC 127
```







Qy	334	PheValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr	353
Db	482	TTCTGTTGGTACAGTTTAAAGATACAGACTCCAGTTTATGCACGGAGAGATGCTTTTACT	541
Qy	354	AspSerAlaIleSerAlaLysValAsnGlyCyluHisLysGlyLysAspLeuGluProTyr	373
Db	542	GACTTCTCTTCACGGCAAGGTGAATGGTGAGCACAGGAGAGACCTGGAGCCCTGG	601
Qy	374	AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsn	393
Db	602	GATCGACGGGAGCTCACGGCCAGCGAGGAGCTGGAG--CTGGAGAAATGATGTCTTAAT	658
Qy	394	GlyTyrAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr	413
Db	659	GGATGGGACCCCAATGATGTTTCGATATAATAGAGAGAAATTATGTGTGGTGTCCACA	718
Qy	414	TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe	433
Db	719	TATGATAGCAGTTTATCTTCATATACCGTTCCTTTAGAAAGGGGCAACTCAGAAGAAATTT	778
Qy	434	LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluLleGluSerSerAlaGlnTyr	453
Db	779	CTTAAACGGGAGGCAAGGCCAAACCACTAGCAGAGAATTTGAATCCAGTGTCTCAGTAC	838
Qy	454	LysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluLysTyrThrAlaVal	473
Db	839	AAACCTCGTGTGCGCCTTTGAAATGATGACCGGAGTGGAGAGAAAAAATACACAGCAGTC	898
Qy	474	GlnArgAsnSerSerGluArgGluGlyHisLeuSerIleAsnThrArgGluAsnLysTyrIle	493
Db	899	CAGAGAAATCTCAGTGACCGGGGGGGGATGGCCCCAACCTAGGGACAATAAATATATT	958
Qy	494	ProProGlyGlnArgAsnArgGluValIleSerTyrGlySerGlyArgGlnAsnSerPro	513
Db	959	CCTCTCGACAAAGAACACAGAGAAGTCTCTCTCTGGGGAGATGGGACACAGAGCTCACCA	1018
Qy	514	ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp	533
Db	1019	CGGATGGCGCAGCCTGGGCCAGGCTCCATGTCGCTCAAGAGCTGTCTTCACACTTCAGAT	1078
Qy	534	PheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTyrProSer	553
Db	1079	TTCAACCCGAAACGCTGGCTCAGACCAAGAGTAGTTAATGAGGTGTTTCCCTGGCCATCG	1138
Qy	554	ProCysProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeu	573
Db	1139	CCTTGGCCCATCTCTCTCTCGCCCACTTCTCGCTACCACTCAGGTCCCAACTCTCTT	1198
Qy	574	ProProArgAlaAlaThrProThrArgPro	583
Db	1199	CCACCTCGGGCAGCAGCCCTTACACGGCCT	1228

## RESULT 10

```

US-08-981-998A-19
; Sequence 19, Application US/08981998A
; Patent No. 684431
; GENERAL INFORMATION:
;
; APPLICANT: PULST, STEFAN M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
;
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: MUEITING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: 55401
; ZIP: 55401
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0. Version #1.30

```

```
Db 362 AAAACATACAGTCCTTAAGTGTGACTTGGTACTTGTGCTGCACATGAGAAAAGTACAGAA 421
Qy SerSerSerGlyProlyshArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333
Db 422 TCCAGTTCCGGGCGCAAAACGTAAGAAATAATGGAGAGTGTGTTTGTTCATAGCTCAGAC 481
Qy 334 PheValValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr 353
Db 482 TTCGTTGTGCTACAGTTTAAAGATACAGACTCCAGTTATGACGGAGAGATGCTTTTACT 541
Qy 354 AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrp 373
Db 542 GACTCTGCTCTCAGCCCAAGGTGAATGTTGAGCACAGAGAGGACCTGGAGCCCTGG 601
Qy 374 AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsn 393
Db 602 GATGCAGGGAGGAGCTACCGCCAGCAGGAGGTGGAG--CTGGAGAAATCATGTGTCTAAT 658
Qy 394 GlyTrpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr 413
Db 659 GGATGGGACCCCAATGACATGTTTCGATATATATGAAGAGAAATTTATGTTGTGTCCACA 718
Qy 414 TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe 433
Db 719 TATGATAGCAGTTTATCTTCATATACGGTTCCTTTAGAAAGGACAACTCAGAGAAATTT 778
Qy 434 LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyr 453
Db 779 CTTAAACGGGAGGCAAGGCAACACCATGTAGCAGAGAAATTTGAATCCAGTGTCTAGTAC 838
Qy 454 LysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluLysTyrThrAlaVal 473
Db 839 AAAGTCGTGTCCCTTCGAGATGATGACCGGAGTGAGGAGAAAGAAATACACAGCAGTC 898
Qy 474 GlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIle 493
Db 899 CAGAGAAACTGCGAGTACCGGGAGGGGCGATGCGCCCAACACTAGGACAAATAATATATT 958
Qy 494 ProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerPro 513
Db 959 CCTCTCGGACAAAGAACAGAGAGTCTTATCTCTGGGGAAGTGGAGACAGAGCTCACCA 1018
Qy 514 ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp 533
Db 1019 CGATGGGCCAGCTGGGCCAGGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGAT 1078
Qy 534 PheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTrpProSer 553
Db 1079 TTCAACCCGACGCTGGCTCAGACCAAGAGTAGTAAATGGAGGTGTTCCCTGGCCATCG 1138
Qy 554 ProCysProSerProSerArgProProProSerArgTyrGlnSerGlyProAsnSerLeu 573
Db 1139 CCTTGCCCATCTCTTCTCTGCCCCACCTTCTCGTACCAGTCAGGTCCCACTCTCTT 1198
Qy 574 ProProArgAlaAlaThrProThrArgPro 583
Db 1199 CCACCTCGGGCAGCCACCCCTACAGGCGCT 1228
```

## RESULT 11

```
US-09-949-016-16240
; Sequence 16240, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16240
; LENGTH: 151088
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151088)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16240
```

```
Alignment Scores:
Pred. No.: 1,366-37 Length: 151088
Score: 1245.50 Matches: 354
Percent Similarity: 49.04% Conservative: 54
Best Local Similarity: 42.55% Mismatches: 176
Query Match: 17.89% Indels: 251
DB: 3 Gaps: 32
```

US-10-802-228-2 (1-1312) x US-09-949-016-16240 (1-151088)

```
Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 2163 ATGCGCTCAGCGCGCGCAGCTCTCGAGTCCCGGTGCCACCGAGTCTCCCGG-TTC 2221
Qy 21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArgSerGlyArg 40
Db 2222 GCCGAGCAGCGTGGCGCGGTGGCGTCCGTCACGCGCGCGCGCGCGCGCGG 2281
Qy 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 2282 GCGCGCGGTGGCGCGCGCGCGTATCCCTCCCGCGCGCGCGCGCGCGCGCC 2341
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 2342 GCGCGCGCGTCCCTCCCGCGCAGAGCTCGCTCCCTCCGCTTCAGACTGTTTGGTAGCAAC 2401
Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro 100
Db 2402 GGCACAGCGCGCGCGCGGTTCGGCCCGGTCCCGCGCGCTCTTGGTCTCGCGCGCCT 2461
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProAlaAla 120
Db 2462 CCGCGCGCTTCTGTCGTCTCTCTCCCTCCGCGCGCGCGCGCGCGCGCGCG 2521
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 2522 CCAACCGCGCGCTCCCGCGTCCCGCGCGGTCCCGCGCGGTTCCTCGCGCTCTCTT 2581
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 2582 GCGCGCGCGGTCTCCCGGTGTCCCGCGCGGTGGAGCGGTATATGGCGCGCTCACC 2641
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 2642 ATGTCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2701
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGly 200
Db 2702 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2761
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerValSer 220
Db 2762 AGCGCGCTTCTACGCTCGCGCGCGCGCGCTTCGCGCGCTCTCTCGCTCTCTCG 2821
Qy 221 SerSerAlaThrAlaProSerSerValAlaAlaThrSerGlyGlyArgProGly 240
Db 2822 TCCTCGGCCACGCTCTCTCTCTCGGTGTCGCGCGCGCGCGCGCGCGCGCGCG 2881
Qy 241 LeuGlyArg-----GlyArgAsnSerAsnLysGlyLeuPro 252
```

```
Db 2882 CTGGGAGGTGGGTGTGCGACACCCAGACCCCTCCGCTCCCGGCCCGCGCTCCCTCCCC 2941
Qy 253 GlnSerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThr 272
Db 2942 CGCGGCCG-----CGCCGCGTCCCGCCCGCGCGTGCAC 2974
Qy 273 SerValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyVal 292
Db 2975 CGC-----CGGGTACCCGGGGTGG- 2995
Qy 293 PheLysThrTyrSerProLysCysAspLeuValLeuAspAlaHisGluLysSerThr 312
Db 2996 -----CTGGGGCCCGCAGCGCGCGGGAGACT 3022
Qy 313 GluSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSer 332
Db 3023 CGCTCGGCGCTCAGCCCGAGGCTCGGCGGTGGCGGCGCGCGGTCTCTC- 3073
Qy 333 AspPheValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPhe 352
Db 3073 ----- 3073
Qy 353 ThrAspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluPro 372
Db 3073 ----- 3073
Qy 373 TrpAspAlaGlyGluLeuThrAlaAsnGluGluLeuAlaLeuGluAsnAspValSer 392
Db 3074 TGGGATTGTCAGGCTGTCCAGCTCCCGCAGCATCCCGCCCTCCCGCGCGGTCAA 3133
Qy 393 AsnGlyTrpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSer 412
Db 3134 GATGCA----- 3139
Qy 413 ThrTyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGlu 432
Db 3140 -----GGGAGCGCGCGCTCCCTCCCTCC- 3163
Qy 453 PheLeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGln 452
Db 3164 -----ACGCGTGTGGAGGGGTCTCGGTAGCGGCGATCGTCAGCCCGCGCTCCCCC 3217
Qy 453 TyrLysAlaArgValAlaLeuGluAsnAspArgSerGluGluGluLysTyrThrAla 472
Db 3218 TTCGCGACGATCTCCGCGCGCAGCGTGGGATGC-TCG----- 3255
Qy 473 ValGlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThr----- 487
Db 3256 -----GGCAGCTCTCCACTCCCGGTTTAGGTGTGAACGTTGGAGGGGTCTGGAGG 3306
Qy 488 -----ArgGluAsnLysTyrIleProProGlyGlnArgAsnArgGluValIleSer 504
Db 3307 CTGTGGTGGCGTTTCCGGAACTGT-CCCCCT-----CCA 3341
Qy 505 TrpGlySer-----GlyArgGln-----AsnSerProArgMetGlyGlnProGly 519
Db 3342 TGGGGGACATCTCTGGAGGGAGAGATTAGGCCCGCGTCCCGTCCCGGTAAAGGGGT 3401
Qy 520 SerGlySerMetProSerArgSerThrSerHisThrSerAspPheAsn-----Pro 536
Db 3402 AGGCACCGGGTCTCTCCGGAATCATCAGGGTCTGTGGGGGTCTCTCCCGCGCCCTCCG 3461
Qy 537 AsnSerGlySerAspGluArgVal--ValAsnGlyGlyValProTrpProSerProCysp 556
Db 3462 AGTCTCGGAAGATTCGGAGGACGGGGTGGAGACAAGTGGGGCTTGGGCC-----C 3512
Qy 556 roSerProSerSer-ArgProProSerArgTyrGlnSerGlyProAsn----- 571
Db 3513 CGCACCCCTCTGCGTTCGTGTCGAGGCGGGCGGGCTCCCGAACTCCCTCGAATC 3572
Qy 572 -----SerLeuProProArgAlaAlaThrProThr----- 581
|||||:|||||
|||:|||||
```

```
Db 3573 GTGGGGCTCCATGTGGCTCCGGCAGCGCTTCCACCC-----TCCCCCACCTGGGGA 3623
Qy 582 -----ArgPro--ProSerArgProProSerArgProSerArg 593
Db 3624 AGGNAAGGGTGGGAGATGCGCGGCCCGCTCCCGCCTTCTCCTTCC-----CCCCCAGA 3680
Qy 594 ProProSerHisProSerAlaHisGlySer-----Pro 604
Db 3681 CCTCTCCGCGCGCGGGTGGTGGCCGATCCCGCATTTGCTGTTCGAGGCCGCGAGTGAGAAG 3740
Qy 605 AlaProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerPro 624
Db 3741 GCGCTGTGGA-----CATCGGTGGGTGAGGGCTGGA-----CCAGCGCTGACCCCT 3788
Qy 625 -----LysAlaGlnArgHis 629
Db 3789 GGAGATCCGGGTGGCGGTGTGGTGGCAGGGGGCGGCACCTCGGCACCTATATCCCAAC 3848
Qy 630 Pro-----ArgAsnHisArgVal-SerAlaGlyArgGlySerIle 642
Db 3849 CCCC GCCCAATTTTCGAAATGCTAGGAGAGAGATTGCAGCAGGGGACGTGGTCGGGT 3908
Qy 642 eSerSerGlyLeuGluPhe-ValSerHisAsn-----ProProSerGluAlaAlaT 659
Db 3909 TCCTGAAGGCAGAAAGCGGGTGTTTACTAGCGTCTTTTTCCTCTCTAAG----- 3958
Qy 659 hrProProVal-----AlaArgThrSerProSer-----GlyGlyThrT 672
Db 3959 --CGGGGTGTAGTAGGGGTGGGGGTCTAGTGTGTCCCGGTAACTGGGTTTGACTC 4016
Qy 672 rpSerSerValSerGlyValPro-----ArgLeuSerProLysThrH 687
Db 4017 GAGGTGTGTGTGTGCAGGAGGCGCTGTGTGGGGTGGCGGGTGTGTGCTCGTATTT 4076
Qy 687 isArgProArgSer-----ProArgGlnAsnSerIleGlyAsnThrP 701
Db 4077 CACGAACCTAAGAAATGCTTAGTGTTCAAAGGGAGAGAAACGTCATAGACTCCATTTC 4136
Qy 701 roSerGlyProValLeuAlaSerProGlnAlaGlyIleIlePro-----ThrGluAlav 719
Db 4137 CATGTGGCCCGTG-----TCTTTAACTTCGGAGTCCCGCAGAGCTTACCAAGGGA 4190
Qy 719 alaAlaMetProIleProAlaAlaSer 727
Db 4191 CGCAAGTCCATTTCCTTGTGCTCA 4216

RESULT 12
US-09-043-303-5
; Sequence 5, Application US/09043303
; Patent No. 6251589
; GENERAL INFORMATION:
; APPLICANT: TSUIJI, Shoji
; APPLICANT: SANPEI, Kazujiro
; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Alaxia Type 2 and
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043,303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (341)..(583)
; FEATURE:
; OTHER INFORMATION: Tsp-2
US-09-043-303-5
```





;; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR  
;; NUMBER OF SEQUENCES: 19  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.  
;; STREET: 119 NORTH FOURTH STREET, SUITE 203  
;; CITY: MINNEAPOLIS  
;; STATE: MINNESOTA  
;; COUNTRY: 55401  
;; ZIP: 55401  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/981,998A  
;; FILING DATE: 11-May-1998  
;; CLASSIFICATION: <Unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: WO 97/42314  
;; FILING DATE: 08-MAY-1997  
;; APPLICATION NUMBER: US 08/727,084  
;; FILING DATE: 08-OCT-1996  
;; APPLICATION NUMBER: US 60/022,207  
;; FILING DATE: 19-JUL-1996  
;; APPLICATION NUMBER: US 60/017,388  
;; FILING DATE: 08-MAY-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: MUETING, ANN M.  
;; REGISTRATION NUMBER: 33,977  
;; REFERENCE/DOCKET NUMBER: 232.00010120  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 612-305-1217  
;; TELEFAX: 612-305-1228  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 516 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; TOPOLOGY: both  
;; MOLECULE TYPE: DNA (genomic)  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-08-981-998A-1

Alignment Scores:  
Pred. No.: 3,328-25 Length: 516  
Score: 857.00 Matches: 166  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.31% Indels: 0  
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-08-981-998A-1 (1-516)

QY 78 GlySerAsnGlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeu 97  
Db 3 GGTAGCAACGGAACGGCGGGCGGCTTCGGCCCGGCTCCCGCGGCTCCTTGGTCTC 62  
QY 98 GlyGlyProProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaPro 117  
Db 63 GCGGGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCC 122  
QY 118 ProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGly 137  
Db 123 CCGGGCGGCGCCACCGGGCGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCC 182  
QY 138 ValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrgly 157  
Db 183 GTCTCTTTGGCGCGCGCGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCC 242  
QY 158 ProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 177

Db 243 CCCCTACCATGTCTGTAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 302  
QY 178 Gln 197  
Db 303 CAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362  
QY 198 ProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSer 217  
Db 363 CCGCGCGCAGCGGCTTCTAGCGTCGCGCGCGCGCGCTTCGCGCTTCGCTCGCTCG 422  
QY 218 ValSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGly 237  
Db 423 GTCTCTCTGCTCTCGGCACGGCTCCCTCTCGGTGGTGGCGGCGACCTCCGCGGCGGG 482  
QY 238 ArgProGlyLeuGlyArg 243  
Db 483 AGGCGCGGCTGGGCAGG 500  
RESULT 15  
US-09-707-919A-13  
; Sequence 13, Application US/09707919A  
; Patent No. 6623927  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: Method for detection of human spinocerebeller ataxia 2  
; FILE REFERENCE: US 443  
; CURRENT APPLICATION NUMBER: US/09/707,919A  
; CURRENT FILING DATE: 2003-02-10  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 459  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Description of  
; OTHER INFORMATION: Artificial derived nucleotide sequence of allelic  
; OTHER INFORMATION: variant of SCA2 gene  
US-09-707-919A-13

Alignment Scores:  
Pred. No.: 1,378-22 Length: 459  
Score: 787.00 Matches: 151  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.34% Mismatches: 0  
Query Match: 11.31% Indels: 0  
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-707-919A-13 (1-459)

QY 72 SerAlaSerAspCysPheGlySerAsnGlyGlyGlyAlaPheArgProGlySer 91  
Db 2 TCCGCTCAGACTGTTTGTAGCAACGGCAACGGCGGCGGCTTCGCGCTTCGCGCTTC 61  
QY 92 ArgArgLeuLeuGlyLeuGlyGlyProProArgProPheValValValLeuLeuProLeu 111  
Db 62 CGGCGGCTCCTTGTCTCGGCGGCTCCCGCGGCTTCGCTCGCTTCCTTCCTCCCTCC 121  
QY 112 AlaSerProGlyAlaProProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAla 131  
Db 122 GCCAGCCGCGGCGGCTCCCGGCGGCGGCTCCCGGCGGCTCCCGGCTCCCGGCTCC 181  
QY 132 SerProProArgSerGlyValSerLeuAlaArgProAlaProGlyCysProArgProAla 151  
Db 182 TCCCGCGGCTTCGCGGCTCTCCTTGGCGGCGGCTCCCGGCTTCGCGCTTCCTCCCG 241  
QY 152 CysGluProValTyrglyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGln 171  
Db 242 TGCAGCGCGGTGTATGGGCGCTTCACCATGTCTCGTGAAGCCCGCAGCAGCAGCAG 301  
QY 172 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 191



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 23, 2005, 06:22:40 ; Search time 1990 Seconds  
(without alignments)  
5451.972 Million cell updates/sec

Title: US-10-802-228-2  
Perfect score: 6961  
Sequence: 1 MSAAPRSPAVATESRRP.....TAHFPYMTSPVQAHHQQL 1312

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10802228/runat\_21122005\_110226\_9755/app\_query.fasta\_1.1479  
-DB=Published Applications\_NA\_Main -QPMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext  
-HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10802228@cgn\_1.1.1549 @runat\_21122005\_110226\_9755 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA\_Main:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	6	US-10-141-541-5
2	6961	100.0	4481	8	US-10-750-323-2
3	6961	100.0	4481	9	US-10-887-553A-493
4	6961	100.0	4481	9	US-10-802-228-1
5	6025.5	86.6	4225	6	US-10-141-541-3
6	2463	35.4	2006	3	US-09-925-298-15
7	2463	35.4	2006	5	US-10-102-806-15

8	1876	27.0	1257	6	US-10-141-541-1	Sequence 1, Appli
9	1876	27.0	1257	8	US-10-750-323-4	Sequence 4, Appli
10	1816.5	26.1	4674	5	US-10-171-581-346	Sequence 346, App
c 11	1042.5	15.0	1707	3	US-09-823-245A-600	Sequence 600, App
c 12	917	13.2	682	9	US-10-956-157-3976	Sequence 3976, Ap
c 13	917	13.2	682	9	US-10-956-157-9211	Sequence 9211, Ap
14	857	12.3	516	8	US-10-750-323-1	Sequence 1, Appli
15	799	11.5	485	3	US-09-918-995-28268	Sequence 28268, A
16	735	10.6	3210	10	US-11-097-143-10706	Sequence 10706, A
c 17	725	10.4	5975	10	US-11-097-143-10705	Sequence 10705, A
c 18	655	9.4	465	4	US-09-925-065A-129085	Sequence 129085,
c 19	598	8.6	493631	5	US-10-087-192-205	Sequence 205, App
c 20	587	8.4	13416	5	US-10-205-032-11	Sequence 11, Appli
c 21	587	8.4	60196	5	US-10-205-032-1	Sequence 1, Appli
c 22	568.5	8.2	64492	7	US-10-378-083-1	Sequence 1, Appli
c 23	568	8.2	18717	6	US-10-156-761-927	Sequence 927, App
c 24	568	8.2	30690	7	US-10-204-862A-1	Sequence 1, Appli
c 25	568	8.2	30690	10	US-11-005-196-1	Sequence 1, Appli
c 26	568	8.2	9025608	6	US-10-156-761-1	Sequence 1, Appli
c 27	566	8.1	11238	5	US-10-205-032-15	Sequence 15, Appli
c 28	564	8.1	88400	8	US-10-844-716-1	Sequence 1, Appli
c 29	556.5	8.0	138203	9	US-10-819-386A-1	Sequence 1, Appli
c 30	543.5	7.8	85915	7	US-10-647-196-1	Sequence 1, Appli
c 31	541.5	7.8	67167	9	US-10-937-730A-3	Sequence 3, Appli
c 32	541	7.8	29467	9	US-10-937-730A-1	Sequence 1, Appli
c 33	528	7.6	65140	7	US-10-203-295-1	Sequence 1, Appli
c 34	528	7.6	125401	7	US-10-203-295-35	Sequence 35, Appli
c 35	522.5	7.5	9025608	6	US-10-156-761-1	Sequence 1, Appli
c 36	520.5	7.5	73995	5	US-10-087-192-208	Sequence 208, App
c 37	520.5	7.5	84428	7	US-10-229-148B-1	Sequence 1, Appli
c 38	515	7.4	22017	8	US-10-211-028-12	Sequence 12, Appli
c 39	515	7.4	90597	8	US-10-211-028-1	Sequence 1, Appli
c 40	511	7.3	20922	3	US-09-942-025-14	Sequence 14, Appli
c 41	511	7.3	67311	3	US-09-942-025-1	Sequence 1, Appli
c 42	501	7.2	34094	3	US-09-963-038A-1	Sequence 1, Appli
c 43	501	7.2	34094	5	US-10-199-550-1	Sequence 1, Appli
c 44	501	7.2	34094	6	US-10-245-603A-1	Sequence 1, Appli
c 45	498.5	7.2	4725	3	US-09-940-316B-24	Sequence 24, Appli

#### ALIGNMENTS

RESULT 1  
US-10-141-541-5  
; Sequence 5, Application US/10141541  
; Publication No. US20030167495A1  
; GENERAL INFORMATION:  
; APPLICANT: Pulst, Stefan M.  
; TITLE OF INVENTION: SCA2 Knockout Animal and Methods of Use  
; FILE REFERENCE: P-CE 5244  
; CURRENT APPLICATION NUMBER: US/10/141,541  
; CURRENT FILING DATE: 2002-05-07  
; PRIOR APPLICATION NUMBER: US 60/289,231  
; PRIOR FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 4481  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (163)...(4101)  
US-10-141-541-5

Alignment Scores:  
Pred. No.: 0  
Score: 6961.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 6  
Length: 4481  
Matches: 1312  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-10-802-228-2 (1-1312) x US-10-141-541-5 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20  
Db 163 ATGCGCTCAGCGCGCGAGCTCTCGAGTCCGCGGTGGCCACCGAGTCTCGCGCTTC 222  
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40  
Db 223 GCCGAGCGCAGGTGGCGCGGTTCGCTCCAGCGCGCGCGCGCGCGAGCGGCGG 282  
QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60  
Db 283 GCGCGGTGGCGCGCGCGCGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCCC 342  
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80  
Db 343 GCGCGCGCTCCCTCCGCGCAGAGCTCGCTCCCTCCGCGCTCAGACTGTTCGTAGCAAC 402  
QY 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgLeuLeuGlyGlyGlyPro 100  
Db 403 GGCACCGCGCGCGCGCGCTTCGCGCCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGGCT 462  
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120  
Db 463 CCGCGCGCTTCGTCTGTCTCTTCCTCCCTTCGCGCGCGCGCGCGCGCTTCGCGCGCG 522  
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140  
Db 523 CCAACCGCGCGCTCCCGCTCGCGCGCGGTGGTTCGCGCGCGTTCGCGCGCTCTCTTG 582  
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160  
Db 583 GCGCGCGCGCTTCGCGGTGTCCCGCGCGCGGTGCGAGCGGTGATGGCGCGCTCACC 642  
QY 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
Db 643 ATGTCTGTGAAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702  
QY 181 GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly 200  
Db 703 CAGCAGCAGCAGCAGCAGCAGCGCGCGCGCGCGCTGCCATGTCCGCAAGCGCGCGC 762  
QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerValSerSer 220  
Db 763 AGCGCGCTTCTAGCTGCGCGCGCGCGCGCTTCGCGGTCTCGCTCTCTCTCTCG 822  
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240  
Db 823 TCCTCGCGCACGGCTCCCTCTCGGTGTGCGCGGACCTCCGCGCGCGGAGGCGCGCGC 882  
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260  
Db 883 CTGGCAGAGGTGGAACAGTAACAAAGAGCTGCCTCAGTCTACGATTTCTTTGATGGA 942  
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280  
Db 943 ATCTATGAAAATATCAGGATGTTTCAATATATACATCAGTTGTGGCTCCAAATGTGAA 1002  
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300  
Db 1003 GTACAGGTGAAAATGAGGATATATATGAGGAGTTTATAAGAGTTTAAACCTTACAGTCCGGAAGTGT 1062  
QY 301 AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg 320  
Db 1063 GATTTGGTACTTGATGCGCCACATGAGAAAAGTACAGAAATCCAGTTCGCGGCGGAAACGT 1122  
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340  
Db 1123 GAAAGAAAATATGGAGAGTATTTTGTTCAAATGTTTCAGACTTTGTGTGTAGTCAAGTTTAAA 1182  
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360  
Db 1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAA 1242

QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380  
Db 1243 GTGAATGGCGAACACAAAGAGAGACCTGGAGCCCTGGGATGCGAGTGAACACTCACAGCC 1302  
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400  
Db 1303 AATGAGAACTTGAGGCTTTGGAAAATGACGCTATCTAATGATGGATGCCAATGATG 1362  
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420  
Db 1363 TTTTCATATAATGAAGAAAATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422  
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440  
Db 1423 TATACAGTGCCTTTAGAAAAGAGATAACTCAGAAGAAATTTTAAACGGGAAGCAAGGCA 1482  
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460  
Db 1483 AACCAAGTTAGCAGAAGAAATTGAGTCAAGTCCCCAGTACAAAGCTCAGTGGCCCTGGA 1542  
QY 461 AsnAspAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480  
Db 1543 AATGATGATAGAGTGAGGAAGAAAATACACAGCTTCAGAGAAATTCAGTGAACGT 1602  
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500  
Db 1603 GAGGGGCACAGCATAAACACTAGGAAAAATAAATAATATATCTCTCTGGACAAAGAAATAGA 1662  
QY 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520  
Db 1663 GAAGTCATATCTCGGGAAGTGGGAGACAGAAATTCACCGGTATCGGCGCAGCTGGATCG 1722  
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540  
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGGTTCA 1782  
QY 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560  
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCTCTCTCT 1842  
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
Db 1843 CGCCACCTTCTCGCTACCACTCAGTCCCAACTCTCTTCACCTCGGCGCACCCCT 1902  
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
Db 1903 ACACGGCGCCCTCCAGGCCCTTCGGGCCATCCAGACCCCTCTCACCCCTCTGCT 1962  
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Db 1963 CATGTTCTCCAGCTCTCTGTCTCTACTATGCTTAAACGATGTCTTCAGAAAGGCCCTCCA 2022  
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Db 2023 AGGATGTCCCAAGGCCCGCAGCATCTCGAAATCACAGAGTTTCTGCTGGAGGGGT 2082  
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
Db 2083 TCCATATCCAGTGGCCCTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCT 2142  
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTyrPheSerSerValSerGlyValPro 680  
Db 2143 CCAGTAGCAAGGACCAAGTCCCTCGGGGGAAACGTGGTTCATCAGTGGTCAGTGGGTCCA 2202  
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
Db 2203 AGATTATCCCCATAAACTCATAGACCCAGGTCTCCACAGACAGAACAGTATTGGAAATACC 2262  
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
Db 2263 CCCAGTGGGCCAGTCTCTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGCC 2322

QY 721 MetProIleProAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
DB 2323 ATGCGCTATTCCAGCTGCATCTCTAGCGCTGTAGTCTGTCATCGAACAGAGCTGTATCC 2382  
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
DB 2383 CTTCTAGTAGGCTAAGATTCAGGCTTCAGATCAGAGGAGCACTCTCTCGAGGG 2442  
QY 761 AsnLysGluAsnLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
DB 2443 AATAAGAAAAATATAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2502  
QY 781 GlyLysSerProValValSerGluHisArgLysGlnLysAspLysLysPheLys 800  
DB 2503 GGTATATCATCCAGTGTCTGTAACATAGAAACAGATTAAGATTAAGATTAAGATTAAG 2562  
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
DB 2563 ATGATTTTAGGTTACAGCAAGTCTCTCTGATCTATGATCAACTACTTAACAA 2622  
QY 821 AsnArgGluGlyLysSerArgAspLeuLysLysLysLysLysLysLysLysLys 840  
DB 2623 AATAGAGAGGAGAAAAATCAAGAGATTTGATCAAGAGCAAAATTAAGATTAAGATTAAG 2682  
QY 841 AspSerPheLysGluAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSer 860  
DB 2683 GATTTCTTTCATTAAGAAATAGCAGCACTGTACCAAGTGGCAGCAGCAAGCGCAATAGC 2742  
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
DB 2743 CCAGCAATTTCCCTTCAATACTTAGTAAACAGAGACACAGAGGGAGCTGAGGTCACT 2802  
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnLysAspLysGluGlu 900  
DB 2803 TCCCAAGGGTTCAGACTTCCAGCCAGCATGTAAACAGAGAAACAGCATTAAGGAAGAG 2862  
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLysAsnProAsnAlaLysGluPhe 920  
DB 2863 AAGAAAGAGCGCAGCTGAGCAAGTATAGAAATCAACATTAAGTCAAGTCAAGTCAAGTTC 2922  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
DB 2923 AACCCAGTTCCTTCTCTAGCCAAAGCTTCTACTACCCAACTTCACTCGGCTCAA 2982  
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrGlnPro 960  
DB 2983 GCACAACTAGCCCATATGATGTTGGTTCATCAACAGCAACTCCAGTTTATCTAGCCT 3042  
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
DB 3043 GTTTGTTTTCACCAATATGATGATCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 3102  
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
DB 3103 TACCAATACCTATGAGCCCAATGAGTGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162  
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020  
DB 3163 AATATGCCCCAACAGCGGCAAGCAGCAGCATCATCAGATGCTGATGATGATGATGATGAT 3222  
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
DB 3223 GCAGCGGCCCCAGGATGAGCCAGCCAGCCAGCTTACTCCAGCAATATGTTGCTAC 3282  
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
DB 3283 AGTCCTCAGAGTTCACCAATCAGCCCTGTTTTCAGCATGTGCAATATCATGTCTCAG 3342  
QY 1061 HisProHisValTyrSerProValLysGlnGlnAlaArgMetMetAlaProProThr 1080  
DB 3343 CATCTCTATGCTATGCTCTGTAATACAGGGTAATGCTAGATGATGATGATGATGATGAT 3402  
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100

DB 3403 CACGCCAGCTGGTTTAGTATCTTTCAGCAACTCAGTACGGGCTCATCAGGAGACG 3462  
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
DB 3463 CATGCGATGTATGATGCTCCCAATTTACCATACAAGGAGACAGCCCTTCTTCTTCTAC 3522  
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
DB 3523 TTGTGCATTTCCAGCGGCTCCCTGCTCAGCAGTATGCGACCTTAAGCTTACCCTGTCAC 3582  
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160  
DB 3583 CCACATATCTCCACACCTCAGCCTTACGCTACCCCACTGGACAGCAGCAAGCAACAT 3642  
QY 1161 GlyLysSerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180  
DB 3643 GGTGAAGTCACTCTGCAACCTCAGTCTGTTTACGACCATCAGCAGCGCCCGCAGGCT 3702  
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTACGCAATTTACCGCGGGCTTGGCCAACT 3762  
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
DB 3763 CCACCTCATGACACCTGCTCCACAGCAGTCCACAGCAAGTATTTCCAGCAGCA 3822  
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
DB 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGTTCAGCGCGGTATACCAACCCACCC 3882  
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
DB 3883 CACATGGGCCCATCTACTCAGGCTCATGTACAGTCCAGGAATGTTCTTCTCATCTCACT 3942  
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
DB 3943 GCCCATGCCCAATGATGCTAATGACGACACAGCAGCCAGCGGCTCCCGAGCGCCCTC 4002  
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300  
DB 4003 GCTCAAGTGCATACAGCCCATTCAGTCTCGACAACAGCGCATTTCCCTTATATGACG 4062  
QY 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312  
DB 4063 CACCCTCTAGTACAAGCCCAACCAACAGCAGTTCG 4098

## RESULT 2

US-10-750-323-2  
; Sequence 2, Application US/10750323  
; Publication NO. US20050032083A1  
; GENERAL INFORMATION:  
; APPLICANT: Pulst, Stefan M  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCERECELLAR  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Mueking, Raasch & Gebhardt, P.A.  
; STREET: 119 North Fourth Street  
; CITY: Minneapolis  
; STATE: Minnesota  
; COUNTRY: USA  
; ZIP: 55401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 30-Dec-2003  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:



```

; APPLICATION NUMBER: US/08/727,084
; FILING DATE: 08-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Muetting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/305-1220
; TELEFAX: 612/305-1228
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: cdna
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 163..4101
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-750-323-2

Alignment Scores:
Pred. No.: 0 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-802-228-2 (1-1312) x US-10-750-323-2 (1-4481)
QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB 163 ATGCGCTCAGCGCGCGAGCTCTCGGAGTCCGCGGTGGCCACCGAGCTCTCGCGCTTC 222
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
DB 223 GCCGACGCGAGGTGGCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGG 282
QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB 283 GCGCGCGGTGGCGCGCGCGCGCGAGCCGTATCCCTCCGCGCGCGCGCGCGCGCGCC 342
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB 343 GCGCGCGCGCGCGCGCGCGCGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro 100
DB 403 GCGAACGCGCGCGCGCGCTTTCGGCGCGCGCTCCGCGCGCGCTCTTGGTCTCGCGCGCGCT 462
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB 463 CCGCGCGCGCTTCGTCTGTCTTCTCCCGCTCCGCGCGCGCGCGCGCGCGCGCGCG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 523 CCNACCGCGCGCTCCCGCTCGCGCGCGCGGTGGCTCCCGCGCGCGGTTCGGCGCTCTCCCTTG 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 583 GCGCGCGCGCGCTCCCGGTGTCCCGCGCGCGGTGCGAGCGGTGTATGGGCGCGCTCACC 642
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
QY 181 GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly 200
DB 703 CAGCAGCAGCAGCAGCAGCAGCGCGCGCGCGGTGCCATGTCCGACGCGCGCGCGCG 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
```

```

DB 763 AGCGCCCTTTAGCGTCCCGCGCGCGCTTCGCGCGCTCTCGTCTCTCGTCTCTCTCG 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240
DB 823 TCCTCGCGCGCGCTCCCTCTCGTGGTCCGCGGACCTCCGCGGCGGAGGCGCGCGCG 882
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrLysSerPheAspGly 260
DB 883 CTGGCAGAGGTGCAAAACAGTAACAAAGGACTGCCTCAGTCTACGATTTCTTTTGATGGA 942
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 943 ATCTATGCAAAATATGAGGATGGTTTATATATGAGGAGGTTTTTAAACTTACAGTCCGAAGTGT 1002
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAAGTGAATAATGGAGGTATATATGAGGAGGTTTTTAAACTTACAGTCCGAAGTGT 1062
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
DB 1063 GATTTGGTACTTGTATGCCGACATGAGAAAGTAGTACAGATCCAGTTCGGGCGCGCAACGT 1122
QY 321 GluGluLeuMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 1123 GAAGAAATAATATGGAGGTATTTTCTTCAAAATGTTTTCAGACTTTGTGTGTACAGTTTAAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAA 1242
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGCGCAACACAAAGAGAGACCTTGGAGCGCTGGAGTGCAGGTGAATCACAGCC 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
DB 1303 AATGAGGAACCTTGAGGCTTTTGGAAAAATCAGCTATCTAATGGATGGGATCCCAATGATATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerSerSer 420
DB 1363 TTTTCGATATATGAAGAAAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTGCCTTTAGAAAGAGATAACTCAGAACAAATTTTAAACGGGAAGCAAGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACCAAGTTAGCAGAAAGAAATTTAGTCAAGTCCCAAGTACAAAGCTCAGAGTGGCGCTGGA 1542
QY 461 AsnAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATAGGAGTGAGGAAGAAAAATACACAGCAGTTTCAGAGAAATTTCCAGTGAACGT 1602
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
DB 1603 GAGGGGCACAGCATAAACACTAGGGAAAAATAAATATATATCTCTCGGACAAAGAAATAGA 1662
QY 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAAGTCATATCTCTGGGAAGTGGGAGACAGAAATTCACCGCTATGGGCGAGCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTCAACCCGAAATCTTGTTTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
DB 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTCCCATCTCTCTCTCTCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
```

Db	1843	CGCCACCTTCTCGCTACAGTCAGTCCCAACTCTCTTCCACCTCGGCGCAGCCACCCCT	1902
Qy	581	ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla	600
Db	1903	ACAGGCGCCCTCCAGGCGCCCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT	1962
Qy	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
Db	1963	CATGGTTCTCCAGCTCCTGTCTCTACTATGCTTAAACGATGCTCTTCAAGAGGGCTCCA	2022
Qy	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
Db	2023	AGGATGTCCTCCAAAGGCCCGCAGCATCTCTCGAATCAAGATTTCTCTGGAGGGGT	2082
Qy	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro	660
Db	2083	TCCATATCCAGTGGCTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCCT	2142
Qy	661	ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro	680
Db	2143	CCAGTAGCAAGGACCACTCCTCGGGGGAAACGTGTCTATCATGTGTCTAGTGGGGTTCCA	2202
Qy	681	ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr	700
Db	2203	AGATTATCCCTTAAACTCATAGACCAGGTCTCCAGACAGAACAGTATTGGAAATACC	2262
Qy	701	ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla	720
Db	2263	CCCAGTGGCCAGTCTCTGCTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGCC	2322
Qy	721	MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr	740
Db	2323	ATGCCTATTTCAGCTGCATCTCTACGCTGTAGTCTCGATCGAACAGAGCTGTACC	2382
Qy	741	ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly	760
Db	2383	CCTTCTAGTGAGGCTTAAGATTCCAGCTTCAAGATCAGAGGCAGAACTCTCTCGAGGG	2442
Qy	761	AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys	780
Db	2443	AATAAGAAATATTAACCCATGAACATCACCTAGCTTCTCAAAAGCTGAACAA	2502
Qy	781	GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLysLysPheLys	800
Db	2503	GGTATATCACAGTTGTTCTGAACATAGAAACAGATTGATGATTTAAGAAATTAAG	2562
Qy	801	AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys	820
Db	2563	AATGATTTTAGGTTACAGCCAAAGTTCTACTTCTGAATCTATGATCAACTACTTAAACAA	2622
Qy	821	AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys	840
Db	2623	AATAGAGAGGAGAAATCAAGAGATTGATCAAGACAAATTAACCAAGTGCTAAG	2682
Qy	841	AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer	860
Db	2683	GATTCCTTCATGAAATATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCGCAATAGC	2742
Qy	861	ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr	880
Db	2743	CCCAGCATTTCCCTCTCAATACTTAGTAACCGGAGCACAAGAGGGACCTGAGGTCACT	2802
Qy	881	SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu	900
Db	2803	TCCCAAGGGTTCCAGCTTCCAGCCCGCATGTATAACAAGAAAGACGATAGGAGAG	2862
Qy	901	LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe	920
Db	2863	AAGAAAGCAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAAGGAGTTC	2922
Qy	921	AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln	940
Db	2923	AAACCAGTTCCTTCTCTAGCCCAAGCCCTTCTACTACCCCAACTTCACTCGGCCCTCAA	2982

Qy	941	AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnPro	960
Db	2983	GCACAACCTAGCCCATCTATGGTGGTTCATCAACAGCAACCTCCAGTTTATACTCAGCCT	3042
Qy	961	ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980
Db	3043	GTTCGTTTGGACCAAAATATGATGTATCCAGTCCAGTCAGCCAGCGGTGCAACCTTTA	3102
Qy	981	TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro	1000
Db	3103	TACCAATACCTATGACGCCCATGTCAGTGAATCAAGCAACACATATATAGACGATACCA	3162
Qy	1001	AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaSerAlaMetMetHisProAlaSer	1020
Db	3163	AATATGCCCCAACACAGCGCAAGACCATCATCAGAGTGCCATGATGACCCAGCGTCA	3222
Qy	1021	AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr	1040
Db	3223	GCAGCGGGCCCACTTGCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTAC	3282
Qy	1041	SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln	1060
Db	3283	AGTCCTCAGCAGTCTCCAAATCAGCCCTTGTTCAGCATGTGCCATATCAGTCTCAG	3342
Qy	1061	HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr	1080
Db	3343	CATCTCATGTCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGCACCAACA	3402
Qy	1081	HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr	1100
Db	3403	CACGCCAGCCTGGTTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACG	3462
Qy	1101	HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr	1120
Db	3463	CATGCGATGTATGTCATGTCCTCCAAATTTACCATAACAACAGGAGACAAGCCCTTCTTCTAC	3522
Qy	1121	PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis	1140
Db	3523	TTTGCCATTTCCAGCGGCTCCCTTGCTCAGCAGTATGCCACCCCTAACGCTACCCCTGCAC	3582
Qy	1141	ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis	1160
Db	3583	CCATATCTCCACACCCCTCAGCCTTCCAGTACCCCACTCCGACAGCAGCAAGCCCAACAT	3642
Qy	1161	GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla	1180
Db	3643	GGTGGAGTCACTCTCCACCCAGTCTGTTTCAGCACCATCAGCACCGCCCGCCAGGCT	3702
Qy	1181	LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr	1200
Db	3703	CTCCATCTGGCCAGTCACAGCAGCAGTCCAGCCATTTACCACCGGGGCTTGGCCCAACT	3762
Qy	1201	ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla	1220
Db	3763	CCACCTCTCATGACACTGCTCCACAGCAGTCGCTCCAGCAATAGTTTCCAGCAGCA	3822
Qy	1221	GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro	1240
Db	3823	CAACAGACTGTCTTTACGATCCATCTCTCAGCTTCCAGCCGGGCTATACCAACCCACCC	3882
Qy	1241	HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr	1260
Db	3883	CACATGGCCCACTACCTCAGGCTCATGATCAGTCAGGAATGGTTCCTTCTCATTCACACT	3942
Qy	1261	AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu	1280
Db	3943	GCCCATGCGCAATGATGCTAATGACACACACCCCGGCTCCCGAGGCGGCCCTC	4002
Qy	1281	AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr	1300
Db	4003	GCTCAAGTGCATACAGCCCATTCAGTCTCCACAACAGCGCATTTCCCTATATGACG	4062

```
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db |||||
4063 CACCCTTCAGTACAAAGCCCAACCAACAGCAGTTG 4098

RESULT 3
US-10-887-553A-493
; Sequence 493, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; with insulin signalling dysregulation
; FILE REFERENCE: 4-32362
; CURRENT APPLICATION NUMBER: US/10/887, 553A
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-493

Alignment Scores:
Pred. No.: 0 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-802-228-2 (1-1312) x US-10-887-553A-493 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db |||||
163 ATGCGCTCAGCGCGCGAGCTCTCGGAGTCCCGGCTGGCCACCGAGTCTCGCGCTTC 222

QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db |||||
223 GCCGAGCAGAGTGCCCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGAGCGGGCGG 282

QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db |||||
283 GCGCGGTGGCGCGCGCGCGGACCGTATCCTCCGCGCGCCCTCCCGCGCGCGCGCC 342

QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db |||||
343 GCGCGCGCTCCCTCCCGCAGAGCTCGCTCCTCCGCTCCAGCTGTGTTGGTAGCAAC 402

QY 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
Db |||||
403 GCAACGCGCGCGCGCGGTTCGCGCGCGCTCCCGCGCGCTCCCTGGTCTCGCGCGGCT 462

QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
Db |||||
463 CCGCGCGCTTCGTGCTGCTCTCTCCCGCTCCGCGCGCGCGCGCGCGCGCGCGCG 522

QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db |||||
523 CCAACCGCGCGCTCCCGCTCGCGCGCGGTGGTCCCGCGCGGTTCGCGCGCTCTCT 582

QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db |||||
583 GCGCGCGCGCTCCCGCTGCTCCCGCGCGGTGGTCCGCGCGGTATGGCGCGCTCACC 642

QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db |||||
643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702

QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
```

```
Db |||||
703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762

QY SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerValSerSer 220
Db |||||
763 AGCGGCTTCAGCGTCCCGCGCGCGCGCTTCGCGCTCTCTGCTCTCTGCTCTCTCG 822

QY SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240
Db |||||
823 TCCTCGGCGCACGGCTCCCTCTCGTGTGCGCGCGACCTCCGCGCGCGCGAGGCCGCG 882

QY LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db |||||
883 CTGGCAGAGGTGCAAAACAGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTATGGA 942

QY IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db |||||
943 ATCTATGCAAAATAGAGATGTTTCAATATCTTACATCAGTTGTGGCTCCAAATGTGAA 1002

QY ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db |||||
1003 GTACAAGTGAAAAATGGAGGTATATATGAAGGAGTTTTTAAACTTTACAGTCCGAAGT 1062

QY AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db |||||
1063 GATTTGGTACTTGATGCGCGCATGAGAAAAAGTACAGAAATCCAGTTCGGGCGCGAAGCT 1122

QY GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db |||||
1123 GAAGAAATATAGGAGATTTTGTTCAAATGTTTCAGACTTTTGTGTGCTACAGTTTAAA 1182

QY AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db |||||
1183 CATATGGACTCCAGTTATGCAAAAGAGATGCTTTACTGACTCTGCTATCAGTCTAAA 1242

QY ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380
Db |||||
1243 GTGAATGGCGAACCAAAAGAGAGAGCCCTGGAGCCCTGGGATGCAGGTGAATCAGACCC 1302

QY AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
Db |||||
1303 AATGAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGATGGGATCCCAATGATATG 1362

QY PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db |||||
1363 TTTTCGATATATGAAGAAAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422

QY TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db |||||
1423 TATACAGTGCCTTTAGAAAGAGATAACTCAGAGAAATTTTTTAAACCGGGAAGCAAGGCA 1482

QY AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db |||||
1483 AACAGTTAGCAGAAAGAAATTTGAGTCAAGTCCCGCAGTACAAAGCTCGAGTGGCCCTGGA 1542

QY AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db |||||
1543 AATGATGATAGAGTGGAGAGAAAAATACACAGCAGTTCCAGAGAAATTTCCAGTCAAGCT 1602

QY GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db |||||
1603 GAGGGGCACAGCATAAACACTAGGAAAAATAAATATATCTCTCTGGGCAAAAGAAATAGA 1662

QY GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db |||||
1663 GAAGTCTATCTCTGGGGAAGTGGAGACAGAAATTTACCGCGGTATGGGCGCAGCTGGATCG 1722

QY GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db |||||
1723 GGCTCCATGCCATCAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTCTGGTTCA 1782

QY AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
Db |||||
```

1783 GACCAAGAGTAGTTAAATGGAGGTGTCTCCCTGGCCATCGCCCTTGCCCATCTCCTCTCT 1842  
Db  
561 ArgProSerArgTyrGlnSerGlyProAsnSerLeuProArgAlaAlaThrPro 580  
Qy  
1843 CGCCACCTCTCTCGTACCAGTCAGTCCCACTCTCTCCACTTCGGGCGCCACCCCT 1902  
Db  
581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600  
Qy  
1903 ACACGGCCGCCCTCCAGGCGCCCTCGCGGCCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
Db  
601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Qy  
1963 CATGGTCTCCAGCTCTGCTCTACTATGCCCTTAACGCGATGCTTTCAGAGGGGCTCCA 2022  
Db  
621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Qy  
2023 AGGATGTCCCAAGGCCCGCCAGCACATCTCGAATCACAGAGTTTCTCTGGAGGGGT 2082  
Db  
641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAlaAlaThrPro 660  
Qy  
2083 TCATATCCAGTGGCCCTAGAAATTTGATGCCACCAACCCAGTGAAGCAGTACTCTCT 2142  
Db  
661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680  
Qy  
2143 CCAGTAGCAAGGACCACTCTCGGGGGGAACGTGTCTCATGTGTCTAGTGGGGTTCCA 2202  
Db  
681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
Qy  
2203 AGATTATCCCTTAAACTCATAGACCAGGTCTCCAGACAGACAGTATGGAAATACC 2262  
Db  
701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
Qy  
2263 CCCAGTGGGCGAGTTCTTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAAGCTGTGCG 2322  
Db  
721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
Qy  
2323 ATGCTTATTCACGTGCATCTCTACGCCCTGTAGTCCGTGCATCGAACAGAGCTGTACC 2382  
Db  
741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
Qy  
2383 CCTCTAGTGAGGCTAAGATTCAGGCTTCAGATTCAGAGGCGAGAACTCTCTCGAGGG 2442  
Db  
761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
Qy  
2443 AATAAGAAATATTAAACCAATGAACATCACTACCTAGCTTCTCAAAAGCTGAAAAACAA 2502  
Db  
781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysPheLys 800  
Qy  
2503 GGTATATCACCGTTGTTCTGAACATAGAAACAGATGATGATTTAAAGAAATTTAAG 2562  
Db  
801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
Qy  
2563 AATGATTTTAGTTACGCCAAGTTCTACTTCTGAATCTATGATCACTACTATAACAA 2622  
Db  
821 AsnArgGluGlyGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
Qy  
2623 AATAGAGAGGAGAAAAATCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTCTAAG 2682  
Db  
841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
Qy  
2683 GATTCTTTCAATTGAAATAGCAGCAGCACTGTATCCAGTGGCAGCAGCAAGCCGAAATAGC 2742  
Db  
861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
Qy  
2743 CCAGAGCATTTCCCTCTCAATCTTAGTAAACGAGACAGAGGGGACCTGAGGTCACT 2802  
Db  
881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGlyLysAspAspLysGluGlu 900  
Qy  
2803 TCCCAAGGGTTCAGACTTCCAGCCCGCAGCATGTAAACAGAGAAAGACGATAAGGAAGAG 2862  
Db  
901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
Qy  
2863 AAGAAGACGACGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAAAGGATTC 2922  
Db

921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
Qy  
2923 AACCCACGCTTCTCTCTCAGCCAAAGGCTTCTACTACCCCAACTTCACCTCGGCTCAA 2982  
Db  
941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
Qy  
2983 GCACAACCTAGCCCATCTATGGTGGGTATCAACAGCCCACTCCAGTTTATCTACTCAGCCT 3042  
Db  
961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
Qy  
3043 GTTGTGTTTGGCACCAATATGATGTATCCAGTCCAGTGAGCCAGCGGTCAACCTTTA 3102  
Db  
981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
Qy  
3103 TACCCAATACCTATGACGCCCATGCGCAGTGAATCAAGCCCAAGACATATAGACAGTACCA 3162  
Db  
1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020  
Qy  
3163 AATATGCCCCCAACGCGGCAAGCAGCATCATCAGATGCCATGATGACCCAGGCTCA 3222  
Db  
1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
Qy  
3223 GCAGCGGGCCCCACCCGATTCGACCCACCCAGCTTACTCCACCAATATGTTGCCCTAC 3282  
Db  
1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
Qy  
3283 AGTCCTCAGCAGTTCCTCAAAATCAGCCCTTGTTCAGCATGTGCCATTTATCAGTCTCAG 3342  
Db  
1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
Qy  
3343 CATCTCTCATGTCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGACCCACCA 3402  
Db  
1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100  
Qy  
3403 CAGGCCACGCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462  
Db  
1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
Qy  
3463 CATGCGATGATGTCATGTCTCAAAATTACCATAACAACAGGAGACAAGCCCTTCTTTCTAC 3522  
Db  
1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
Qy  
3523 TTTGGCCATTTCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCCCTAACGCTACCTCTCAC 3582  
Db  
1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160  
Qy  
3583 CCACATATCTCCACCCCTCAGCTTCCAGTACCCCTTGGACAGCAGCAAGCAACAT 3642  
Db  
1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180  
Qy  
3643 GGTGGAAGTCATCTGACCCAGTCTGTGTCAGCACCATCAGCACCGCCGCGCCAGGCT 3702  
Db  
1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
Qy  
3703 CTCCATCTGGCAGTCCACAGCAGCAGTCCAGCATTTACACGCGGGCTTGGCCAACT 3762  
Db  
1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
Qy  
3763 CCACCTCCATGACACTGCTTACGATCCATCTCTCAGCTTCAGCGCGGTATACCAACCCACC 3822  
Db  
1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
Qy  
3823 CAACAGACTGTCTTTACGATCCATCTCTCAGCTTCAGCGCGGTATACCAACCCACC 3882  
Db  
1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
Qy  
3883 CACATGGCCACGCTACCTCAGGCTCATGATCAGTTCAGGAATGGTCTCTCTCTCACTCAACT 3942  
Db  
1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280  
Qy  
3943 GCCCATGCGCAATGATGCTTAATGACGACACAGCCACCCGCGGTCTCCAGGCGCGCTC 4002  
Db



Db 1783 GACCAAGAGTAGTAAATGAGAGGTTCCTCCCTGGCCATCGCCTTGCCCAATCTCCTTCCTCT 1842  
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580  
Db 1843 CGCCCACTTCCTCGCTACCAAGTCAGTCCCAACTCTCTCCACCTCGGCGACCCACT 1902  
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600  
Db 1903 ACACGGCGCCCTCCAGGCGCCCTCGGCGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962  
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620  
Db 1963 CATGGTTCCTCCAGCTCTCTCTACTATGCTTAACCGCATGCTCTCAGAGGGGCTCCA 2022  
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640  
Db 2023 AGGATGTCCTCCCAAGGCCCGCAGCATCTCTGAAATCACAGAGTTTCTGCTGGGAGGGT 2082  
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660  
Db 2083 TCCATATCCAGTGGCTTAGAATTTGTATCCCAACACCCAGTGAAGCAGTACTCTCT 2142  
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680  
Db 2143 CCAGTAGCAAGGACGAGTCCCTCGGCGGGAACGTGTCTATCAGTGTCTCAGTGGGTTTCA 2202  
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700  
Db 2203 AGATTATCCCTTAAACTCATAGACCCAGTCTCCAGACAGAACAGTATTGGNAATACC 2262  
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720  
Db 2263 CCCAGTGGCCAGTCTCTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGC 2322  
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740  
Db 2323 ATGCTCTATTCCAGCTGCATCTCTAGCTGTAGTCTGTAGTCTGCAATGCAAGAGCTGTATCC 2382  
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760  
Db 2383 CTTCTAGTAGGCTTAAAGATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCGAGGG 2442  
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780  
Db 2443 AATAAGAAATATTAATTAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAATAACAA 2502  
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspleuLysLysPheLys 800  
Db 2503 GGTATATCACCAAGTCTTTCTGAACATGAAGAAACAGATTGATGATTAAAGAAATTTAAG 2562  
Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820  
Db 2563 ATGATTTTAGTTACAGCAAGTCTTCTCTGATCTATGATCAACTACTAACAACAA 2622  
Qy 821 AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840  
Db 2623 AATAGAGGGGAGAAATCAAGAGATTGATCAAGACAAATTAAGCAAGTGTCTAAG 2682  
Qy 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860  
Db 2683 GATTCTTTTATTGAAATAGCAGCAACTGTACCAGTGGCAGCAGCAAGCGCAATAGC 2742  
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880  
Db 2743 CCAGCATTTCCCTTCAATCTATTAGTAACGAGACACAAAGGGGAGCTGAGGTCACT 2802  
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900  
Db 2803 TCCCAAGGGGTTCCAGACTTCCAGCCCGAGCATGTAAACAGAGAAACACGATGAAGGAAG 2862  
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
Db 2863 AAGAAAGACGCACTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAGGAGTTTC 2922

Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
Db 2923 AACCCAGTTCCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACTCGGCTCAA 2982  
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960  
Db 2983 GCACAACTTACCCATCTATGGTGGTTCATCAACAGCCAACTCCAGTTTATCTCAGCT 3042  
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980  
Db 3043 GTTTGTTTTCACCAATATATGATGATCCAGTCCAGTGAGCCCGGCGTGAACCTTTA 3102  
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000  
Db 3103 TACCAATATCCATATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCA 3162  
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020  
Db 3163 AATATGCCCCAACAGCGGCAAGACAGCATCATCAGAGTGCCATGATGCCACCGCTCA 3222  
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040  
Db 3223 GCAGCGGCGCCACCGATTGCGAGCCACCCAGCTTACTCCAGCAATATGTTGCCTAC 3282  
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060  
Db 3283 AGTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAG 3342  
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080  
Db 3343 CATCTCATGTCTATAGTCTCTGTAATACAGGTAATGCTAGAAATGATGCCACCAACA 3402  
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyValHisGluGlnThr 1100  
Db 3403 CACGCCACGCTGTTTGTAGTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3462  
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120  
Db 3463 CATGGATGATGATGCTGCTCCCAATTTACATACAAAGGAGACAGCCCTTCTTTCTAC 3522  
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnTyrAlaHisProAsnAlaThrLeuHis 1140  
Db 3523 TTTGCTCATTTCCAGGGCTCTCTGCTCAGCAGTATGCGACCCCTTAACGCTACCTGTGAC 3582  
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160  
Db 3583 CCACATATCTCCACACCTCAGCTTACCTACCCCTGAGACAGCAGCAAGCCCAACAT 3642  
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180  
Db 3643 GGTGAAGTCTATCTCTGCCACCATCTGTTTACGACCATCAGCAGCAGCCCGCCAGGCT 3702  
Qy 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200  
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCATTTTACCGGGGCTTGGCCCACT 3762  
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220  
Db 3763 CCACCTCTCAATGACACTGCTCCCAACAGCAGTCCGACACAGAAATAGTTTCCAGCAGCA 3822  
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240  
Db 3823 CAACAGACTGTCTTTACGATCTCTCTCTCAGCTTACGCGCGGTATACCAACCCACCC 3882  
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260  
Db 3883 CACATGGGCCGCTACCTCAGGCTCATGTACAGTCAAGAAATGTTCTTCTCATCAACT 3942  
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaLeu 1280  
Db 3943 GCCCATGCGCCCAATGATGTAATGACGACAGCAGCCCGCGGTCCCCAGCGCCCTTC 4002





Db 1422 GGCACAAAGAAACAGAGAAGTCTCTATCTCGGGGAAGTGGAGACAGAGCTCCACCGGATG 1481  
QY 516 GlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAspPheAsn 535  
Db 1482 GGCACAGCTGGGCCAGGCTCCATGCGGTCAAGAGCTGTTCTCACACTTCAGATTCAAC 1541  
QY 536 ProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTrpProSerProCys 555  
Db 1542 CCGAAGCTGGCTCAGACCAAGAGTAGTTAATGAGAGTGTTCCTCGGCCATCGCCTTGC 1601  
QY 556 ProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuProPro 575  
Db 1602 CCATCTCATCTCTCGCCACCTTCTCGCTACCAAGTCAGGTCCCAACTCTCTTCCACCT 1661  
QY 576 ArgAlaAlaThrProThrArgProProSerArgProProSerArgProProPro 595  
Db 1662 CCGGCGAGCCACCACCATACAGCGCGCCCTCCAGGCCCTCCAGGCCATCCAGACCCCG 1721  
QY 596 SerHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMetSer 615  
Db 1722 TCTCACCCCTCTGCTCATGTTCTCCAGCTCCTGTCTCTACTATGCTTAAACGCGATGCT 1781  
QY 616 SerGluGlyProProArgMetSerProLysAlaGlnArgHisProArgHisArgVal 635  
Db 1782 TCAGAAAGGACCCCAAGAGTGTCTCCAAAGGCACAGCGCCACCTCGGAATCACAGAGTC 1841  
QY 636 SerAlaGlyArgGlySerLeuSerSerGlyLeuGluPheValSerHisAsnProProSer 655  
Db 1842 TCTGCTGGGAGAGGCTCCATGTTCTAGTGGCTTAGAATTTGTATCCCAATATCCCCCAAGT 1901  
QY 656 GluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerSerVal 675  
Db 1902 GAAGCAGCTGCTCTCCAGTGGCAAGACCAAGTCTCGAGGGGAGGAGTGGTCTCAGTG 1961  
QY 676 ValSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsn 695  
Db 1962 GTCAGTGGGGTTCCAAGGTATCTCCCAAACTCACAGACCCCAAGGTCTCCAGGCAGAGC 2021  
QY 696 SerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleIlePro 715  
Db 2022 AGCATTTGGAAACTCTCCAGCGGGCTGTGCTTGTCTCTCCCAAGCTGCAATCATCCCT 2081  
QY 716 ThrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSer 735  
Db 2082 GCAGAGCCGTTTCCATGCTGTTCCCGCGCATCTCCGACTCTCGCAGCCCTGCATCC 2141  
QY 736 AsnArgAlaValThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGln 755  
Db 2142 AACAGAGCACTGACCCCATCTATTGAGGCAAAAGATTCCAGGCTTCAAGATCAGAGGCAG 2201  
QY 756 AsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSer 775  
Db 2202 AACTCTCTCGCGGAGTAAGAAAAATGTTAAAGCAAGTGAAACATCACCTAGCTTTTCA 2261  
QY 776 LysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAsp 795  
Db 2262 AAAGCTGACAAACAAAGTAGTGTCCACAGTGTGTTCTGAAACACAGAAAAACAGATTGATGAC 2321  
QY 796 LeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAsp 815  
Db 2322 TTTAAGAAGTTTAAGATGATTTTAGGTTACAGCCAAAGCTTACATCTGAATCTATGGAT 2381  
QY 816 GlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIle 835  
Db 2382 CAACTACTAAGCAAAAAATAGAGAGGAGAAAGTCAACGAGATTGATTAAAGATAAAACG 2441  
QY 836 GluProSerAlaLysAspSerPheIleGlu-----AsnSerSerSerAsnCysThrSer 853  
Db 2442 GAAGCAAGTGTAAAGATAGTTTTCATTGACAGCAGCAGCAGCAGCAACTGTACCAAGT 2501  
QY 854 GlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThrGluHis 873

Db 2502 GGCAGCAGCAAGACCAACAGCCCTAGCATCTCCCTTCCATGCTTAGTAATGACAGCAC 2561  
QY 874 LysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCysLysGln 893  
Db 2562 AAGAGGGGCGCTGAGGTACATCCCAAGGGTGCAGACTTCCAGCCGACCTTGCACAAACAA 2621  
QY 894 GluLysAspAspLysGluGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeu 913  
Db 2622 GAGAAGGATGACAGAGAGAGAAAGAACACACACAGAGCAGGTTAGGAAATCGACATTG 2681  
QY 914 AsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLysProSerThrThr 933  
Db 2682 AATCCCATGCAAGAGAGTTCAACCTCTGTTCTCTCAGCCAAAGCCTTCTACTACC 2741  
QY 934 ProThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnPro 953  
Db 2742 CCAACGTCACCTCGGCCCTCAAGCACCAACCCAGCCCATCTATGTTGGTTCATCAGCAGCCA 2801  
QY 954 ThrProValThrThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal 973  
Db 2802 GCTCCAGTGTACACTCAGCCCTGTGTGCTTCGCACCAATATATGATATCCCTGCCAGTG 2861  
QY 974 SerProGlyValGlnProLeuTyrProIleProMetThrProMetProValAsnGlnAla 993  
Db 2862 AGCCCGGGGCTACAACTTTATACCAATACCTATGACGCCCATGCTGTGAACCAAGCC 2921  
QY 994 LysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAspGlnHisHis 1011  
Db 2922 AAGACATATAGCAGAGGTAAAGTACCAATATGCCCCAACAGCAGCAGCAACCAATCAT 2981  
QY 1012 GlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThrProPro 1031  
Db 2982 CAAAGCACCATATGACCCAGCCCTCCCGGAGGCGCCACCCATCTGTAGCCACCCGCC 3041  
QY 1032 AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGlnProLeuVal 1051  
Db 3042 GCTTACTCCATCAGTACGTTGCTTACAGCCCTCAGCAGTTCCTCCCAATCAGCCTTTGGTC 3101  
QY 1052 GlnHisValProHisThrGlnSerGlnHisProHisValTyrSerProValIleGlnGly 1071  
Db 3102 CAGCATGTGCCCATTTATCAGTCTCAGCATCTCTCATGTGTACAGTCTCTCATACAAGGT 3161  
QY 1072 AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAla 1091  
Db 3162 AATGCCAGGATGATGCGCACCCAGCAGCATGTCTAGCCCTGGTTAGTGTCTTCTTCAGCT 3221  
QY 1092 ThrGlnTyrGlyValHisGluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyr 1111  
Db 3222 GCTCAGTTCGGGGCTCAGCAGCAGCAGCAGCATGTATGATGTCTCCCAATTTACCATAC 3281  
QY 1112 AsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGln 1131  
Db 3282 AACAGGAGACAAAGCCCTTCTTCTTACTTTGCCATTTCCACCGCTCTCTCGCTCAGCAG 3341  
QY 1132 TyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThr 1151  
Db 3342 TATGCACATCTTAATCCGCCCTTGCAATCCACATACCTCCCATCTCTCAGCCTTCGGCCACT 3401  
QY 1152 ProThrGlyGlnGlnGlnSerGlnHisGlySerHisProAlaProSerProValGln 1171  
Db 3402 CCCACCGGACAGCAGCAAGCCAGCATGTTGGAAGTCACCTCGACCCAGTCTCTGTTTCAG 3461  
QY 1172 HisHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGlnGlnGlnSerAla 1191  
Db 3462 CACCATCAGCAGCAGGCTGCCAGGCTCTTCATCTGCCAGTCCACAGCAGCAGTCCGCC 3521  
QY 1192 IleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSerAsnThrGln 1211  
Db 3522 ATTTATCATCGGGGCTGGCACCAACCACTTCCATGACACTCTGCTCTTAATACACAG 3581  
QY 1212 SerProGlnAsnSerPheProAlaAlaGlnThrValPheThrIleHisProSerHis 1231  
Db 3582 TCTCCAGAGCAGGAGTTTCCAGCAGCAGCAACACAGACAGTCTTCCACCATCCACCTTCTCAT 3641

```
QY 1232 ValGlnProAlaTyrThrAsnProProHisMetAlaHisValProGlnAlaHisValGln 1251
Db 3642 GTTCAGCGGGCATACACACCCACCCACATGGGCCAGGTACCTCAGGCTCATGTACAG 3701
QY 1252 SerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMetThrThrGln 1271
Db 3702 TCAGAAATGGTTCCTCTCATCAACTGCCATGGCCCAATGATGCTAATGACGACACAG 3761
QY 1272 ProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProIleProValSer 1291
Db 3762 CCACCC---GGTCCCAAGGCCGCTCGCTCAAGTGCACTACAGCCCATTCAGTTTCG 3818
QY 1292 ThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaHisHisGlnGlnGln 1311
Db 3819 ACAACAGCGCAATTCCTTATATGACGACGCCCTTCAGTACAAGGCCACCCACACAGCAG 3878
QY 1312 Leu 1312
Db 3879 TTG 3881

RESULT 6
; Sequence 15, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2006
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (862)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (1006)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-298-15

Alignment Scores:
Pred. No.: 1,42e-138 Length: 2006
Score: 2463.00 Matches: 460
Percent Similarity: 98.72% Conservative: 1
Best local Similarity: 99.50% Mismatches: 2
Query Match: 35.38% Indels: 4
DB: 3 Gaps: 1

US-10-802-228-2 (1-1312) x US-09-925-298-15 (1-2006)
QY 850 AenCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSer 869
Db 6 AGCTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCAATACTTAGT 65
QY 870 AenThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerPro 889
Db 66 AACCGGAGCACAAGAGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACTTCCAGCCCA 125
QY 890 AlaCysLysGlnGluLysAspAspLysGluGluLysLysAspAlaAlaGluGlnValArg 909
Db 126 GCATGTAAACAGACAGAAAGACGATTAAGGAGAGAAAGACCGCACTTGACCAAGTTAGG 185
QY 910 LysSerThrLeuAsnProAsnAlaAlaLysGluPheAsnProArgSerPheSerGlnProLys 929
```

```
Db 186 AAATCAACATTGAATCCCAATGCAAGGAGTTCACCCACGCTTCTCTCTCAGCCAAAG 245
QY 930 ProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetValGly 949
Db 246 CTTTCTACTACCCCAACTTCACCTCGGCCTCAAGCACAACTAGCCCACTATATGGTGGGT 305
QY 950 HisGlnGlnProThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTyr 969
Db 306 CATCAACAGCCCACTCCAGTTTATACACAGCTGTTGTTTGGACCAAAATATGATGAT 365
QY 970 ProValProValSerProGlyValGlnProLeuTyrProIleProMetThrProMetPro 989
Db 366 CCAGTCCCAAGTGGAGCCAGGCGTCAACCTTTATACCAATACCTATGACGCCCATGCCA 425
QY 990 ValAsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGln 1007
Db 426 GTGAATCAAGCCCAAGACATATAGACGAGGTAAAGTACCAAAATATGCCCAACAGCGGCA 485
QY 1008 AspGlnHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAla 1027
Db 486 GACCAGCATCATCAGAGTGCCATGATGCACCCAGCGTCAGCAGCGGCCCCACCGATTGCA 545
QY 1028 AlaThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsn 1047
Db 546 GCMACCCCAACAGCTTACTCCACGCAATATGTCCTACAGTCTCCTCAGCAGTTCCCAAT 605
QY 1048 GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerPro 1067
Db 606 CAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCTCATGTCTATAGTCT 665
QY 1068 ValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal 1087
Db 666 GTAATCAGGGTAATGTAGTAATGATGGCACCAACACACAGCCCGCCAGCTGGTTAGTA 725
QY 1088 SerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysPro 1107
Db 726 TCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGATGCGATGATGATGTCCTCC 785
QY 1108 LysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySer 1127
Db 786 AAATTACCATACAAAGGAGACAAAGCCCTTCTTCTACTTTGGCATTTCCACGGGCTCC 845
QY 1128 LeuAlaGlnGlnTyr-AlaHisProAsnAlaThrLeuHisProHisThrProHisProGln 1147
Db 846 CTTGCTCAGCAGTAGTATGNGCRACCTTACGCTACCTGCACCCACATACTCCACACCTCA 905
QY 1147 nProSerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaPr 1167
Db 906 GCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCAACATGGTGAAGTCATCTCTGCACC 965
QY 1167 oSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla-LeuHisLeuAlaSerProG 1187
Db 966 CAGTCTCTGTTCAGCACCATCAGCACCAGCGCCGAGGCTTCTCATCTGGCCCATGCCAC 1025
QY 1187 GlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProA 1207
Db 1026 AGCAGCAGTCAAGCATTTTACCACCGGGGCTTGGCCCAACTCCACCTCCATGACACCTG 1085
QY 1207 laSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrI 1227
Db 1086 CCTCCAAACAGCAGTCGCGCAGAAATAGTTTCCCAAGCAGCAGCAACAGACTGTGTTTACGA 1145
QY 1227 leHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisValProG 1247
Db 1146 TCCATCTCTTCTCAGCTTTCAGCGCGGTATACCAACCCACCCACATGGCCCACTTACCTC 1204
QY 1247 InAlaHisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetL 1267
Db 1206 AGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAACTGCCCATGGCCCAATGATGC 1265
QY 1267 euMetThrThrGlnProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnP 1287
```

Db 1266 TAATGACGACACAGCCACCCGGGGTCCCGAGCGCCCTCGCTCAAAAGTGCACTACAGC 1325  
QY 1287 roileProValSerThrAlaHisPheProTyrMetThrHisProSerValGlnAlaH 1307  
Db 1326 CCATTCCAGTCTCGACAAACAGCGGCATTTCCCTATATGACGACCCCTTCAGTACAAGCCC 1385  
QY 1307 isHisGlnGlnGlnLeu 1312  
Db 1386 ACCACCAACAGCAGTTG 1402

RESULT 7  
US-10-102-806-15  
; Sequence 15, Application US/10102806  
; Publication No. US20030054421A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103PIC1  
; CURRENT APPLICATION NUMBER: US/10/102,806  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 2006  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (862)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (1006)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-102-806-15

Alignment Scores:  
Pred. No.: 1,42e-138 Length: 2006  
Score: 2463.00 Matches: 460  
Percent Similarity: 98.72% Conservative: 1  
Best Local Similarity: 98.50% Mismatches: 2  
Query Match: 35.38% Indels: 4  
DB: 5 Gaps: 1

US-10-802-228-2 (1-1312) x US-10-102-806-15 (1-2006)

QY 850 AsnCysThrSerGlySerIysProAsnSerProSerIleSerProSerIleLeuSer 869  
Db 6 AGCTGTACAGTGGCAGCAGCAGCGAATAGCCCGCAGCATTTCCCTTCATACTACT 65  
QY 870 AsnThrGluHisIysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerPro 889  
Db 66 AACACGAGCACAAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCCAGACTTCCAGCCCA 125  
QY 890 AlaCysGlnGlnIysAspAspIysGluIuLysAspAlaGluGlnValArg 909  
Db 126 GCATGTAAACAGAGAAAGACGATAGGAAGAGAGAAAGACGACCTGAGCAAGTTAGG 185  
QY 910 LysSerThrLeuAsnProAlaLysGluPheAsnProArgSerPheSerGlnProLys 929  
Db 186 AAATCAACATTGATCCCATGCAAGAGGTTCAACCCAGTTCTCTCTCAGCCCAAG 245  
QY 930 ProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetValGly 949  
Db 246 CCTTCTACTACCCCACTTCACCTCGCGCTCAAGCAACACCTAGCCCATCTATGTGGGT 305  
QY 950 HisGlnGlnProThrProValThrGlnProValCysPheAlaProAsnMetMetTyr 969

Db 306 CATCAACAGCCAACTCCAGTTTATATCTACGCTGTTTGTGTTTGCACCAAAATATGATGAT 365  
QY 970 ProValProValSerProGlyValGlnProLeuTyrProIleProMetThrProMetPro 989  
Db 366 CCAGTCCAGTGGAGCCAGCGGTGCAACCTTTATACCAATACCTTATACGCCCATGCA 425  
QY 990 ValAsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGln 1007  
Db 426 GTGAATCAAGCCAGACATATAGAGCAGGTAAGTACCAAAATATGCCCAACAGCGCAA 485  
QY 1008 AspGlnHisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAla 1027  
Db 486 GACCAGCATCATCAGATGTCATGATGACCCAGCGTCAGCAGCGCGGCCACCGATTGCA 545  
QY 1028 AlaThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsn 1047  
Db 546 GMACCCCCAGCAGCTTACTCCACGCAATATGTTCCTCAGCTCCTCAGCAGTTCCCAAT 605  
QY 1048 GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerPro 1067  
Db 606 CAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAGCATCTCATGTCTATAGTCT 665  
QY 1068 ValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal 1087  
Db 666 GTAATACAGGGTAATGCTAGAAATGATGACCAACACACAGCCAGCCTGTTTAGTA 725  
QY 1088 SerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysPro 1107  
Db 726 TCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACGATGCGATGATCATGTCTCC 785  
QY 1108 LysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySer 1127  
Db 786 AAATTACCATACCAAGAGAGACAGCCCTCTTCTTACTTTGCCATTTCCACGGGCTCC 845  
QY 1128 LeuAlaGlnGlnTyr-AlaHisProAsnAlaThrLeuHisProHisThrProHisProG 1147  
Db 846 CTTGCTCAGCAGTATGNGCRACCTTAACGCTACCTGACCCACACACTACTCCACACCTCA 905  
QY 1147 nProSerAlaThrProThrGlyGlnGlnSerGlnHisGlySerHisProAlaPro 1167  
Db 906 GCCTTCAGCTACCCCACTGGAGCAGCAAGCAAGCAACATGTTGGAAGTCATCTGCACC 965  
QY 1167 oSerProValGlnHisHisGlnAlaGlnAlaGlnAla-LeuHisLeuAlaSerProG 1187  
Db 966 CAGTCTGTTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1025  
QY 1187 InGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProA 1207  
Db 1026 AGCAGCAGTCAGCCATTTACCACGCGGGCTTGGCCCAACTCCACCTCCATGACACCTG 1085  
QY 1207 laSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrI 1227  
Db 1086 CCTCCAACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1145  
QY 1227 leHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisValProG 1247  
Db 1146 TCCATCTCTTCACGTTACGCGGGGTATACCAACCCACCCACATGGCCCGCAGTACCTC 1205  
QY 1247 InAlaHisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetL 1267  
Db 1206 AGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACTGCCCATGCGCAATGATGC 1265  
QY 1267 euMetThrThrGlnProProGlyGlyProGlnAlaLeuAlaGlnSerAlaLeuGlnP 1287  
Db 1266 TAATGACGACACAGCCACCCGGGTCGCCAGCGCCCTCGCTCAAGTGCACTACAGC 1325  
QY 1287 rolleProValSerThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaH 1307  
Db 1326 CCATTCCAGTCTCGACAAACAGCGGCATTTCCCTATATGACGACCCCTTCAGTACAAGCCC 1385  
QY 1307 isHisGlnGlnLeu 1312  
Db 1386 ACCACCAACAGCAGTTG 1402

```

RESULT 8
US-10-141-541-1
; Sequence 1, Application US/10141541
; Publication No. US20030167495A1
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.
; TITLE OF INVENTION: SCA2 Knockout Animal and Methods of Use
; FILE REFERENCE: P-CE 5244
; CURRENT APPLICATION NUMBER: US/10/141,541
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/289,231
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1255)
US-10-141-541-1
Alignment Scores:
Pred. No.: 1,55e-103 Length: 1257
Score: 1876.00 Matches: 365
Percent Similarity: 89.77% Conservative: 21
Best Local Similarity: 84.88% Mismatches: 18
Query Match: 26.95% Indels: 26
DB: 6 Gaps: 5

US-10-802-228-2 (1-1312) x US-10-141-541-1 (1-1257)
QY 157 GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGln 176
Db 8 GGGCGCTCACCATCTCGTGAAGCCGCGCGCAG----- 43
QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db 44 -----CCGCCGCGCGCCGCGCGC 67
QY 197 LysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaProSerProSerSer 216
Db 68 AAGCCCGCGCGC---GGCCTGCTCTCGTCCGCGCGCGCGCGCGC---TCGCGCGCG 121
QY 217 SerValSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSer 234
Db 122 GTGACCTCGGCTTCGCTGGTGGCGCGCGCGCGCGCGCGTCTTCTCGCGCGCGCC 181
QY 235 ---GlyGlyGlyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
Db 192 GCGGCGCGCGCGCTCCCGGCTGGCGAGAGTGGAAACATAGGAAAGGACTGCTCTCAG 241
QY 254 SerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSer 273
Db 242 CCTACGATTTCTTTTGATGTAATCTATGCAACGGTGGAGTGGTTTCATATATTAGGTCA 301
QY 274 ValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPhe 293
Db 302 GTTGTGTGATCGAATGTAAGTACAGTGAAGTGAACCGGAGCATATATGAAGAGTGT 361
QY 294 LysThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGlu 313
Db 362 AAACATACAGTCTTAAGTGTGACTTGTGTTGATGTCGACATGAGAAAAGTACAGAA 421
QY 314 SerSerSerGlyProLysArgGluIleMetGluSerIleLeuPheLysCysSerAsp 333
Db 422 TCCAGTTCCGGGCCAAAACGTGAAGAAAATAATGGAGAGTGTGTTTGTTCATAATGCTCAGAC 481
QY 334 PheValValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr 353
Db 482 TTCGTTGTGGTACAGTTTAAAGATACACAGACTCCAGTTTATGACCGGAGAGATGCTTTTACT 541

```

```

QY 354 AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrp 373
Db 542 GACTCTGCTCTCAGCGCAAGGTGAATGGTGAGCACAAAGGAGGACCTCGAGCCCTGG 601
QY 374 AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAspValSerAsn 393
Db 602 GATCAGGGGAGCTCACGGCCAGCGAGGAGCTGGAG---CTGGAGAATGATGTGCTTAAT 658
QY 394 GlyTrpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr 413
Db 659 GGATGGGACCCCAATGACATGTTTCGATATAATGAAGAGAATTATGGTGGTGCTCCACA 718
QY 414 TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe 433
Db 719 TATGATAGCAGTTTATCTTCATATACGGTTCTTTAGAAAGGGCAACTCAGAAGAATTT 778
QY 434 LeuLysArgGluAlaArgAlaAsnGlnIleuAlaGluIleGluSerSerAlaGlnTyr 453
Db 779 CTTAAACGGGAGGCAAGGCAAAACAGTTAGCAGAGAATTTGAATCCAGTGCTCAGTAC 838
QY 454 LysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluLysTyrThrAlaVal 473
Db 839 AAGCTCGTGTGCGCCCTTGAGAAATGATGACCGGAGTGAGGAAGAAAATAACACAGCAGTC 898
QY 474 GlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIle 493
Db 899 CAGAGAAATGCACTGACCGGAGGGGCATGGCCCAACACTAGGGACAAATAATATATT 958
QY 494 ProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerPro 513
Db 959 CCTCTGGACAAAGAAACAGAGAAGTCTATCTCTGGGGAAGTGGGACAGAGCTCACCA 1018
QY 514 ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp 533
Db 1019 CGGATGGCCAGCCTGGCCAGGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGAT 1078
QY 534 PheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTrpProSer 553
Db 1079 TTCAACCGGAGCGTGGCTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCG 1138
QY 554 ProCysProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeu 573
Db 1139 CTTTGGCCATCTCCTCTCTCGCCACCTTCTCGTACCAGTCCAGTCCAGTCCCAACTCTCT 1198
QY 574 ProProArgAlaAlaThrProThrArgPro 583
Db 1199 CCACCTCGGCGAGCCACCCCTACACGGCCT 1228

RESULT 9
US-10-750-323-4
; Sequence 4, Application US/10750323
; Publication No. US20050032083A1
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCERE BELLIAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Muelling, Raasch & Gebhardt, P. A.
; STREET: 119 North Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/750,323
; FILING DATE: 30-Dec-2003

```

```

RESULT 10
US-10-171-581-346
; Sequence 346, Application US/10171581
; Publication No. US2003010426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 346
; LENGTH: 4674
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF034373
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-346

```

## Alignment Scores:

Pred. No.: 2,16e-99 Length: 4674  
Score: 1816.50 Matches: 504  
Percent Similarity: 48.22% Conservative: 174  
Best Local Similarity: 35.85% Mismatches: 381  
Query Match: 26.10% Indels: 350  
DB: 5 Gaps: 52

US-10-802-228-2 (1-1312) x US-10-171-581-346 (1-4674)

```
QY 19 ArgPheAlaAlaAlaAArgTProGlyTrpArgSerLeuGlnArgProAlaArgSer 38
DB 6 CGCTCTCCTCCTCCTAGATGGATGGCGCGCGGAGAGGCGCTCCTCATTTCTTAGATG 65
QY 39 GlyArgGlyGlyGlyAlaAlaProGlyProGlyProSerAlaAlaProProPro 58
DB 66 GGATGGCGCGCGCGAGAGATGCTCTCCTCCTCCTCCTCCTCCTCCTCCT 125
QY 59 GlyProGlyProProProSerArgGlnSer-----SerProProSerAlaSerAspCys 76
DB 126 CGGCTCGCGCGCGCTTCCTCGCGCGCGGCTTCTCTCTCTCCTCCTCCTCCTCCTCCT 170
QY 77 PheGlySerAsnGlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGly 96
DB 171 -----CCGACACCGCG----- 182
QY 97 LeuGlyGlyProProArgProPheValValLeuLeuLeuLeuLeuLeuLeuLeuLeu 116
DB 183 -----GGCTCCCGCGCGCG-----CCGACGCGCGCGCGCGCGCTC 218
QY 117 ProProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSer 136
DB 219 CGATCCCGCTCGCTTCCCGCGC-TCTCCAGCGCGCGCGCGCGCGCGCGCTCTCTCC 277
QY 137 GlyValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyr 156
DB 278 CTCCTCTCTCTA-ATTCCCTCTCCGACGCT----- 309
QY 157 GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGln 176
DB 310 -----GCCATCATGTTGAAGCCTCAGCGCTTACACAGCCCTCCCGCGCGCGCGCG 360
QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
DB 361 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 196 GLYsProGlyGly-----SerGlyLeuLeu 204
DB 421 CGGCTCTCGCGCGCGCTGCGCACCTCTGGGCTCTCTCGCGCGCGCTCCAGCGCGCGCT 480
QY 204 uAlaSerProAlaAlaAla-----ProSerProSerSerSerSerValse 219
DB 481 CCCCTGCTCGGGCTGTGGCGCTGCGGAGCGGCTCCCGCGGAGCGCGAGCGCA 540
QY 219 rSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgPr 239
DB 541 TCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 239 o-GlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheA 259
DB 601 TCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
QY 259 spGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysC 279
DB 658 AAGCGCTTACAAATCCAGATGCTGCTATTTCTTACAGCTGTGTGGGCTCCACTT 717
QY 279 ysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProL 299
DB 718 GTGATGTAAGGTGAAATGATGATACCTTATGAGGGTATCTTCAAGACGCTAAGCTCAA 777
QY 299 ysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrCluSerSerGlyProL 319
DB 778 AGTTGAAGTACCGGTGATGCTGTGACCGGAAGACATCTGAGCCAGCGGTGCGCTCCT 837
```







```
Db 471 CGGCTGGGAGTGGCCAACTCTTCCAGCTCCATCCCTGTGACCTCATCAGTCTCAGATCC 412
QY 727 -----SerProThrProAlaSerProAlaSerProAlaSerProAlaValThrProSerSe 743
Db 411 TGGAGTGGGCTCCATTTCTCCAGTCTCTCCAAAGATCTCCCTGGCCCCCACA----- 360
QY 743 rGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysG1 763
Db 359 -----GATGTAAAGAACTCTTACCAAGAAACCTGGGAG 325
QY 763 uAenlleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSe 783
Db 324 AACTCTGGAGCCCGAGAGCTG-----GCTCGATAGCTGGGAAAT 283
QY 783 rProValValSerGluHisArgLys-----GlnlleAspAspLeuLysLysPheLysAs 801
Db 282 CCCTGGTCTTCAGATGAACAGAAACGATTCCAACTGGAAGAACTGAGAAAGTTTGGGCG 223
QY 801 nAspPheArgLeuGlnProSerSerThrSerGlu---SerMetAspGlnLeuLeuAsnLys 820
Db 222 CCAGTTTAAGCTTCAGCCAGTAGTCCCTCCCTGAGAACAGCCTGGATCTCTTCTCCCTCC 163
QY 820 sAenArgGluGlyGluLysSerArgAspLeuLysAspLysLysLysLysLysLysLysLys 840
Db 162 G-----ATCTTAAAGGAG-----GAGCCCAAGGAAA 136
QY 840 sAspSerPheIleGluAsn-----SerSerSerAs 850
Db 135 GGAGAAAGAGTTGATGGTCTGTGACTTCAGAGCCCATGGGCTCTCCCTCTCCCA 76
QY 850 nCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAs 870
Db 75 GACAGAGTCCGTATCGATGAAGAGGAGGACAAACCCCTGGCACCATCA-----GGAGG 22
QY 870 nThrGluHisLysArgProGlu 878
Db 21 CACTGAG-----GGGCCAGAG 6

RESULT 12
US-10-956-157-3976/c
; Sequence 3976, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3976
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-3976

Alignment Scores:
Pred. No.: 3,17e-46 Length: 682
Score: 917.00 Matches: 170
Percent Similarity: 90.53% Conservative: 2
Best Local Similarity: 89.47% Mismatches: 0
Query Match: 13.17% Indels: 18
DB: 9 Gaps: 1

US-10-802-228-2 (1-1312) x US-10-956-157-3976 (1-682)

QY 1058 GlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAla 1077
Db 682 AAGTCTCAGCATCTCATGCTATAGTCTGTATACAGGGTAATGCTAGAATGATGGCA 623
QY 1078 ProProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHis 1097
```

```
Db 622 CCACCAACACACGCCCGAGCTGTAGTATCTTCTCAGCAACTCAGTACGGGCTCAT 563
QY 1098 GluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerPro 1117
Db 562 GAGCAGACGATCGATGTAT----- 542
QY 1118 SerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAla 1137
Db 541 -----GTTTCCACGGGCTCCCTGTGCTCAGCAGTATGCGCACCTTAACGCT 497
QY 1138 ThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGln 1157
Db 496 ACCTTGACCCACATATCTCCACACCTCAGCTTACCTACCCCTCGACAGCAGCAA 437
QY 1158 SerGlnHisGlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAla 1177
Db 436 AGCCAAACATGGTGAAGTCTCTGCGCAGTCCACAGCAGCTGAGCCATTTACCACCGGGGCTT 377
QY 1178 AlaGlnAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeu 1197
Db 376 GCGCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCTGCTTACGATCCATCTCTCAGCTT 317
QY 1198 AlaProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhe 1217
Db 316 GCGCCAACTCCACCTCCATGACCTGCTCAACACGACGCTCGCCACAGATAGTTTC 257
QY 1218 ProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThr 1237
Db 256 CCAGCAGACACAACAGACAGTCTTTACGATCCATCTCTTCTCAGCTTCCAGCCGCGGTATACC 197
QY 1238 AsnProProHisMetAlaHisValProGln 1247
Db 196 AACCCACCCACATGGCCCGACGCTACCTCAG 167

RESULT 13
US-10-956-157-9211
; Sequence 9211, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-9211

Alignment Scores:
Pred. No.: 3,17e-46 Length: 682
Score: 917.00 Matches: 170
Percent Similarity: 90.53% Conservative: 2
Best Local Similarity: 89.47% Mismatches: 0
Query Match: 13.17% Indels: 18
DB: 9 Gaps: 1

US-10-802-228-2 (1-1312) x US-10-956-157-9211 (1-682)

QY 1058 GlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAla 1077
Db 1 AAGTCTCAGCATCTCATGCTATAGTCTGTATACAGGGTAATGCTAGAATGATGGCA 60
QY 1078 ProProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHis 1097
Db 61 CCACCAACACACGCCCGAGCTGTAGTATCTTCTCAGCAACTCAGTACGGGCTCAT 120
```

```
QY 1098 GluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerPro 1117
Db 121 GAGCAGACGCATGCCATGAT----- 141
QY 1118 SerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAla 1137
Db 142 -----GTTTCCACGGGCTCCCTTGTCTCAGCAGTATGCCACCCCTAACGCT 186
QY 1138 ThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGln 1157
Db 187 ACCCTGCACCCACATACTCCACACCTCAGCCTTCAGCTACCCCACTGGACAGAGCAA 246
QY 1158 SerGlnHisGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAla 1177
Db 247 AGCCAACATGGTGGAGATCATCTGCACCCAGTCTCTTTCAGCACCATCAGCACCCAGGCC 306
QY 1178 AlaGlnAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeu 1197
Db 307 GCCCAGGCTCTCCATCTGCCCAAGTCCACAGCAGCAGTCCAGCAATTTACACGCGGGGCTT 366
QY 1198 AlaProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhe 1217
Db 367 GGGCCAACTCCACCTCCATGACACCTGCCTCCAACACGAGTCCGCCACAGATAGTTTC 426
QY 1218 ProAlaAlaGlnGlnThrValPheThrLeHisProSerHisValGlnProAlaTyrThr 1237
Db 427 CCAGCAGCACACACACTGTCTTTACGATCCATCTCTCAGTTCACGTCGCGCGGTATACC 486
QY 1238 AsnProProHisMetAlaHisValProGln 1247
Db 487 AACCCACCCACATGGCCACGATACCTCAG 516
RESULT 14
US-10-750-323-1
; Sequence 1, Application US/10750323
; Publication No. US20050032083A1
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBILLAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Mueeting, Raasch & Gebhardt, P.A.
; STREET: 119 North Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/750,323
; FILING DATE: 30-Dec-2003
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,084
; FILING DATE: 08-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueeting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/305-1220
; TELEFAX: 612/305-1228
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
```

```
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-750-323-1
Alignment Scores:
Pred. No.: 9,61e-43 Length: 516
Score: 857.00 Matches: 166
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.31% Indels: 0
DB: 8 Gaps: 0
US-10-802-228-2 (1-1312) x US-10-750-323-1 (1-516)
QY 78 GlySerAsnGlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeu 97
Db 3 GGTAGCAACGGAACAGCGGGCGGCGGTTTCGGCCCGGCTCCCGGGCGCTCTTGGTCTC 62
QY 98 GlyGlyProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaPro 117
Db 63 GCGGGGCTCCCGCCCTTTCGTCTCTTCTCCCTCCGCGAGCCCGGGCGCCCT 122
QY 118 ProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGly 137
Db 123 CCGCCCGCGCCCAACCCCGGCTCCCGGCTCGGCGCGGCTCGGCTCCCGCGCGCTTCCGCG 182
QY 138 ValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGly 157
Db 183 GTCTCTTGGCGCGCGGCTCCCGGCTGTCCTCCCGCGGCTCGGAGCCGTTGATGGG 242
QY 158 ProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 177
Db 243 CCCCTCACCATGCTGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 302
QY 178 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
Db 303 CAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362
QY 198 ProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSer 217
Db 363 CCGCGGGCAGCGGCTTCTAGCGTCCCGCGCGCGGCTTCCGCGCTCTCTCTCTCG 422
QY 218 ValSerSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGly 237
Db 423 GTCTCTCTCGTCTCGGCGCAGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
QY 238 ArgProGlyLeuGlyArg 243
Db 483 AGGCGCGCGCTGGGCAGG 500
RESULT 15
US-09-918-995-28268
; Sequence 28268, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28268
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(485)
; OTHER INFORMATION: n = A,T,C or G
```

US-09-918-995-28268

## Alignment Scores:

Pred. No.:	2,74e-39	Length:	485
Score:	799.00	Matches:	145
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.48%	Indels:	0
DB:	3	Gaps:	0

US-10-802-228-2 (1-1312) x US-09-918-995-28268 (1-485)

QY	1128	LeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGln	1147
DB	51	CTTGCTCAGCAGTATGCGACCCCTAACGCTACCTGCACCCACATCTCCACACCTCAG	110
QY	1148	ProSerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaPro	1167
DB	111	CCTTCAGCTACCCCACTGGACAGCAGCAAGCCACATGGTGAAGTCTCTCGCACCC	170
QY	1168	SerProValGlnHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGln	1187
DB	171	AGTCCTGTTTCAGCACCATCAGCACAGCGCGCCAGGCTCTCCATCTGGCCAGTCCACAG	230
QY	1188	GlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProSerMetThrProAla	1207
DB	231	CAGCAGTCAGCCATTTACCAACGCGGGGCTTGGCCCAACTCCACCTCCATGACACCTGCC	290
QY	1208	SerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIle	1227
DB	291	TCCAACACGCGAGTCGCCACAGAAATAGTTTCCACGACAGCAACAGACTGTCTTTACGATC	350
QY	1228	HisProSerHisValGlnProAlaTyrThrAsnProHisMetAlaHisValProGln	1247
DB	351	CATCCTTCTCACGTTTCAGCGGGGTATACCAACCCACATGCCCCCAGGTACCTCAG	410
QY	1248	AlaHisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeu	1267
DB	411	GCTCATGTACAGTCAGGAATGGTTCTTCTCTCATCCAATGCCCATGGCCAAATGATGCTA	470
QY	1268	MetThrThrGlnPro	1272
DB	471	ATGACGACACAGCCN	485

Search completed: December 23, 2005, 14:14:21  
Job time : 2122 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 23, 2005, 06:53:47 ; Search time 320 Seconds  
(without alignments)  
2127.715 Million cell updates/sec

Title: US-10-802-228-2  
Perfect score: 6961  
Sequence: 1 MRSAAAPRSPAVATESRRF.....TAHFYMTHPVSQAHQOOL 1312

Scoring table:

BLOSUM62			
Xgapop 10.0 , Xgapext 0.5			
Ygapop 10.0 , Ygapext 0.5			
Fgapop 6.0 , Fgapext 7.0			
Delop 6.0 , Delext 7.0			

Searched: 4168288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=slp  
-Q=/cgn2/USPRO.spool\_p/US10802228/runat\_21122005\_110226\_9796/app\_query.fasta.1.1479  
-DB=Published Applications NA\_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10802228@cgn 1.184 @runat\_21122005\_110226\_9796  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA\_New:

- 1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US12\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US13\_NEW\_PUB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US14\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	553.5	8.0	116856	7 US-11-143-980-1	Sequence 1, Appli
C 2	461.5	6.6	5682	7 US-11-059-982-2	Sequence 2, Appli
C 3	440	6.3	15720	7 US-11-108-172-1058	Sequence 1058, Ap
C 4	437	6.3	10968	7 US-11-075-185-35	Sequence 35, Appli
C 5	437	6.3	78869	7 US-11-075-185-1	Sequence 1, Appli
C 6	435	6.2	6210	7 US-11-056-470-1	Sequence 1, Appli
C 7	416	6.0	10524	7 US-11-075-185-37	Sequence 37, Appli
C 8	407	5.8	164810	7 US-11-121-086-4	Sequence 4, Appli

C 9	399	5.7	11070	7 US-11-075-185-34	Sequence 34, Appli
C 10	395.5	5.7	153376	7 US-11-121-086-5	Sequence 5, Appli
C 11	395.5	5.7	172543	7 US-11-121-086-6	Sequence 6, Appli
C 12	394.5	5.7	7006	6 US-10-821-234-218	Sequence 218, App
C 13	386	5.5	8730	7 US-11-087-100-1	Sequence 1, Appli
C 14	386	5.5	8730	7 US-11-087-084-1	Sequence 1, Appli
C 15	386	5.5	8730	7 US-11-087-085-1	Sequence 1, Appli
C 16	385	5.5	6360	7 US-11-056-470-2	Sequence 2, Appli
C 17	375	5.4	3408	6 US-10-858-730-40	Sequence 40, Appli
C 18	374	5.4	6683	6 US-10-995-561-473	Sequence 473, App
C 19	374	5.4	6700	6 US-10-995-561-471	Sequence 472, App
C 20	374	5.4	6833	6 US-10-995-561-471	Sequence 471, App
C 21	374	5.4	6871	6 US-10-995-561-474	Sequence 474, App
C 22	370	5.3	6941	6 US-10-432-483-49	Sequence 49, Appli
C 23	364.5	5.2	4598	6 US-10-955-054-71	Sequence 71, Appli
C 24	361.5	5.2	5679	7 US-11-075-185-36	Sequence 36, Appli
C 25	361.5	5.2	6012	6 US-10-467-657-83	Sequence 83, Appli
C 26	361.5	5.2	6012	6 US-10-467-657-6321	Sequence 6321, Ap
C 27	359	5.2	6763	7 US-11-080-991-55	Sequence 55, Appli
C 28	357	5.1	11736	7 US-11-000-463-218	Sequence 218, App
C 29	356.5	5.1	8651	6 US-10-432-483-48	Sequence 48, Appli
C 30	355	5.1	116856	7 US-11-143-980-1	Sequence 1, Appli
C 31	352	5.1	8651	6 US-10-432-483-48	Sequence 48, Appli
C 32	348.5	5.0	4509	7 US-11-087-100-5	Sequence 5, Appli
C 33	348.5	5.0	4509	7 US-11-087-084-5	Sequence 5, Appli
C 34	348.5	5.0	4509	7 US-11-087-085-5	Sequence 5, Appli
C 35	348.5	5.0	6786	7 US-11-069-834-59	Sequence 59, Appli
C 36	347.5	5.0	63984	7 US-11-121-086-26	Sequence 26, Appli
C 37	346.5	5.0	23907	7 US-11-186-731-6	Sequence 6, Appli
C 38	346.5	5.0	24120	7 US-11-186-731-4	Sequence 4, Appli
C 39	341.5	4.9	13650	6 US-10-995-561-155	Sequence 155, App
C 40	341.5	4.9	13711	6 US-10-995-561-154	Sequence 154, App
C 41	341.5	4.9	14194	6 US-10-995-561-153	Sequence 153, App
C 42	340	4.9	14172	7 US-11-075-185-2	Sequence 2, Appli
C 43	339.5	4.9	5913	6 US-10-821-234-789	Sequence 789, App
C 44	339.5	4.9	6360	7 US-11-056-470-2	Sequence 2, Appli
C 45	339.5	4.9	6377	7 US-11-069-834-57	Sequence 57, Appli

ALIGNMENTS

RESULT 1  
US-11-143-980-1/c  
; Sequence 1, Application US/11143980  
; Publication No. US20050272133A1  
; GENERAL INFORMATION:  
; APPLICANT: He, Min  
; APPLICANT: Hucul, John  
; APPLICANT: Haltli, Bradley A.  
; APPLICANT: Wagenaar, Melissa M.  
; APPLICANT: Graziani, Edmund  
; APPLICANT: Summers, Mia  
; APPLICANT: Kulowski, Kerry  
; APPLICANT: Pong, Kevin  
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex  
; FILE OF INVENTION: Polyketide  
; FILE REFERENCE: AM-101426US  
; CURRENT APPLICATION NUMBER: US/11/143,980  
; PRIOR FILING DATE: 2005-06-03  
; PRIOR FILING DATE: 2005-03-23  
; PRIOR FILING DATE: 2005-03-23  
; PRIOR FILING DATE: 2004-06-03  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 116856  
; TYPE: DNA  
; ORGANISM: Streptomyces sp.  
US-11-143-980-1

Alignment Scores: 1.13e-11 Length: 116856  
Pred. No.: 116856





Db 72877 CGCTCAGAGCTCAGCGACTCCAAACACACACCGTTCATCCAGACACACCGCGCGCATGCACACC 72818  
Qy 588 o-----SerArgProSerArgProProSerHisProSer-----AlaHisG1 602  
Db 72817 ACACGCAATCCCAATCCCTCCAAACACTCGGCCACCGCAGCCCGATCCGCCACATCA 72758  
Qy 602 ySerProAlaProValSerThrMetPro-----LysArgMetSerSerGluG1 618  
Db 72757 CAGCGCACCGCCGACCGCAACACCCCAACCCCTCCAACTCCGCGCAGCAGCAGCTGCTCCGCA 72698  
Qy 618 yProProArgMetSerPro-----LysAlaGlnArgHisProArgAsnHisArgVa 635  
Db 72697 TCCGGCGCAGCATCCCGCGAGACTCGCGCAGCAGCAACAGTCTCCACACCGACACCGGCC 72638  
Qy 635 lSerAlaGly-----ArgGlySerIleSerSerGlyLe 646  
Db 72637 AACCCCGCGCCACATGACCCACCGCAGCAGCCGATACCGCGCTCACCAACCGCGCCCT 72578  
Qy 646 uGluPheValSerHisAsnProSerGluAlaAlaThr-ProProValAlaArgThrS 666  
Db 72577 CGACC--CGGCCACACACCGGCCACCGCATCCCGAATCCCGAACCAGCGCCGACACCGCGACCA 72520  
Qy 666 ex-----ProSerGlyGlyThrProSerValVal----- 676  
Db 72519 GCGGAGCAGCCCGCCACCGCCACCGAACCGCAACCTGGTCTCTCCCGGTACCCCCAC 72460  
Qy 676 ----- 676  
Db 72459 CCNACACCGCGCACAACCGCCCGCTCTCATCCACCACCGCCGACATCGATCA 72400  
Qy 677 -----SerGlyValProArg-----LeuSerProLysThrHisArgProA 690  
Db 72399 GACCAACCCCAACGATCCGGATGCTCCAAAGCGATCACCAGCGGCGCAAGACCCACACCCCG 72340  
Qy 690 rGSerProArgGlnAsn----- 695  
Db 72339 ACTGCCCCGGATCCACCATCATCACCACCGACCGCAGCAGCAGCCCGAGTCACACACC 72280  
Qy 696 -----SerI 697  
Db 72279 ACNACCGCGCATCGGCCCGCGCTCCCGAGTCCCTGCACCATGACACCGTCCGCTCGA 72220  
Qy 697 leGlyAsnThrProSerGlyProValLeuAla---SerProGlnAlaGlyIleIleProT 716  
Db 72219 CCGGCAACACGACACACACCGCGCAGCGACCTCGCCAGCTCGCGCATGTGCACAA 72160  
Qy 716 hrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSerA 736  
Db 72159 CTGCGCATCGGCCCGCGCAGCGCCGACGACGCGGACCATCGTCTCCGCGCTCGCCCA 72100  
Qy 736 snArgAlaValThrPro-----SerSerGluAla----- 745  
Db 72099 CCACGACCGCTGCGCGCAGACTCGGCTGCTCGGAGCGGTCCGATCGGGAGCAGT 72040  
Qy 746 -----LysAspSer-----ArgLeuGlnAspG 753  
Db 72039 CGAGGCGGTACCGCCAGACTCGACGAGTACCTCTCTGTCGACGCGCGCGCAGGACG 71980  
Qy 753 lnArgGlnAsnSerProAla-----GlyAsnLysGluAsnIleLysProAsnGluThrS 771  
Db 71979 AGAGGCGGCGAGTGGCGCGCTGAGCGGCTGTTCGGCATCGATACCGAAGAACCGAGAT 71920  
Qy 771 exProSerPheSerLysAlaGluAsnLysGlyIleSerProValVal----- 786  
Db 71919 CCCACGCTCAACCGCATCCAGAACCGCGGTCGCGCGCTCGTCTGGTGTGCTCAG 71860  
Qy 787 -----SerGluHisArgLysGlnIleAspLeuLysPheLysAsnAspP 803  
Db 71859 GGGAGCGCTCAACCGACAGCGCTCAG-CTGGAAGGCATACGTTCG----- 71814  
Qy 803 heArgLeuGlnProSerSerThrSerGlu-----SerMetAsp----- 815  
Db 71813 --CAAATCAACCGCGCAGCAGCCCGGGAACACCGCATCCCGATCCCGATCTGGACACCCCGG 71756

Qy 816 -----GlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeu-----lleLysA 833  
Db 71755 ACGTACAGCCCAACCAAGCGCCATAACCGGTTCAGAGGCTCCGAGCATCCCGACGCAAG 71696  
Qy 833 spLys---lle-GluProSerAlaLysAspSerPheIleGluAsnSerSerSerAsnCys 851  
Db 71695 ACAGGACACCATCCCATCCGCAAG-----CGGTCAACGTC 71657  
Qy 852 ThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLysSerAsnThr 871  
Db 71656 CCGTCAGGCCCCCACTCCAGGAACCAACCAACCGAGCTACCGAGCGCTCTCCAGCCCA 71597  
Qy 872 GluHisLysArgGlyProGluValThr-----SerGlnGlyValGlnThrSer 887  
Db 71596 TCCGGAATCGGACGCTCTCCGCACTGACGACCCCAATACTCCGGCGAGCAGAGCTCT 71537  
Qy 888 SerProAlaCysLysGlnGluLysAspAspLysGluLysLysAspAlaAlaGluGln 907  
Db 71536 TCACCGGCCACCGCAGCGACAGCTTCGACACCGGCGACGCTCGGCGACCGGAATCC 71477  
Qy 908 ValArgLysSerThrLeuAsnProAsnAlaLysGluPhe----- 920  
Db 71476 ACGCCCGGAAGCACC---TCACCGAATCCGCGAGCATCCCGTCCATACGAGCCGAATGA 71420  
Qy 921 -----AsnProArgSerPheSerGlnProLysProSerThrThr-ProTh 935  
Db 71419 AACGCGATGACTCACCGCAACCGCGGTCCGACCCCAACCGCCACCAACATCCACA 71360  
Qy 935 rSerProArgProGlnAlaGlnPro-----SerProSerMetValGlyHisGlnGlnPr 953  
Db 71359 ACGGCTCAACCGCATCTCTATCCCGGAGACACCGCAGCTCGGAGCATTCACCGCC 71300  
Qy 953 oThrProValThrGlnProValCysPheAlaProAsnMetMetTy-ProValProVa 973  
Db 71299 GCNACCCACACCTCCGCGCAGCGCGCGACCTCACCTCCGACCGCGCCACCGCA 71240  
Qy 973 l-----SerProGlyValGlnProLeuTyProIleProMetThrProMetProVa 990  
Db 71239 ACCATCGCACCGCGCAGCGCAACGCTGCATCAAAAGCGCGCGCCGACACCGCCCTC 71180  
Qy 990 lAsnGlnAlaLysThrTyArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHi 1010  
Db 71179 CCGCATCTCCAAACGACACACCGCCCATACGCGCGCACTCCGCGCACT----- 71131  
Qy 1010 shHisGlnSer-----AlaMetMetHisPro-----AlaSe 1020  
Db 71130 -CACCATCGATGACCCAGCAGATAATCCCGCTCACACCCCGACGAGCGATCAGCGA 71072  
Qy 1020 rAlaAlaGlyProProIleAlaAlaThrProAla----- 1032  
Db 71071 TACAGCGACACCTCCAGCGGAACACCGCGCTGCGCATACACCGTCTCCCGACCAAC 71012  
Qy 1033 -----Ty-SerThrGlnTyValAlaTySerProGlnGlnPheProAsnGlnProL 1050  
Db 71011 CCGCATCATACCCCGACACCTCCCGCGCGCTATCCAGATACCGATCCAGTCC 70952  
Qy 1050 euValGlnHisValProHisTyGlnSerGlnHisProHisValTySerProValIleG 1070  
Db 70951 CCGCACCTCATCCACAC---CGCAGCGCAACAC-----CGGATACCTCTCATAC 70904  
Qy 1070 lngLysAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerS 1090  
Db 70903 AA-----CCACGACCATCCCAACCGCT-CGACGCCCTGACC 70866  
Qy 1090 erAlaThrGlnTyGlyAlaHisGluGln-----ThrH 1101  
Db 70865 CGCAAAACACACACCAACCCACCGCAACCGGAACACCCACCGACCGAACCCCAACCC 70806  
Qy 1101 isAlaMetTyAlaCysProLysLeuProTyArgAsnLysGluThrSerSerPheTyP 1121  
Db 70805 AC-----CCAGCAACTCTCTCCAGTCCGACCGACCGACCGAC----- 70765





```
Db 3325 CCTGAATGTGGCCAAAGCCGCTTCCTCCGGGCCAGGGAAGCCCTCCGGGCTGCAGTATG 3384
Qy 753 lNArgGlnAsnSerProAlaGlyAsnLysGluAsnLysProAsnGluThrSer 771
Db 3385 AGAGCAAACTGAGTGGCTGAAGAGCCCCACACGCTTCAGGCCCAAGGAAGCCTGTT 3444
Qy 771 ----- 771
Db 3445 TCCTGGAGCATTTGCACAAACACACAGGGCTCGTCTCGCACCCGCACTACAAGACGGCT 3504
Qy 772 --ProSerPheSerLysAlaGluAsnLys-----GlyIleS 783
Db 3505 TCCCTCTCTTTGAGACGCCCTGCATCGCTCTCTCCCTACCATGCTTACCAGGCGGCC 3564
Qy 783 erProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAsp 803
Db 3565 TCCCTCTCCCACTGACTACCAAAA---GTGACAGAG-----GAGT 3603
Qy 803 heArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArg- 822
Db 3604 TTGACACGCTCTCCACGACGCTGCTGAAACAGCACCCAGCCATGCTCAATAAATATCGC 3663
Qy 823 ----GluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysA 841
Db 3664 TCTCTCTCTCGAGAGAGTCCCG-----AGGTTGAGCCCTCAGCGGA-G 3707
Qy 841 epSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSerP 861
Db 3708 ATGGT-----AATGATCGACCGAATGTTTCATTGAGGAGGA-AGACCACCTTGC 3757
Qy 861 roSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrS 881
Db 3758 CTTGGATA-----AACAGCTGGCCAAGGAGAAGC----- 3786
Qy 881 erGlnGlyValGlnThrSerSerProAlaCysLysGlnLysAspAspLysGluGlu 901
Db 3786 ----- 3786
Qy 901 ysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPheA 921
Db 3787 -----CGACAGATGACTGTTCTTCTCCCTCGCTCGCTCGGCT 3823
Qy 921 snProArgSerPheSerGlnProLysProSerThr-----ThrProThrSerProA 938
Db 3824 CCCCCA-----TCGACAGCCTCTTCGAGGGTCATCGGCTTCCCGGCACCGGCCCT 3874
Qy 938 rgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyr 958
Db 3875 GTCGTCTTCAGTCCCGGGGCT-----CCACCCAGCCCTCCACCA 3916
Qy 958 hrGlnProValCysPhe-----AlaProAsnMetMetTyrProV 971
Db 3917 CTTGCCCAACCAAGTTGTATCGGCACGGGGGCGAGGGCTCCCTTCGTCACCTG 3976
Qy 971 alPro---ValSerProGlyValGlnProLeuTyrProIleProMetThrProMetProV 990
Db 3977 GSCCGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4036
Qy 990 alAsnGlnAlaLysThrTyr-ArgAlaValProAsnMetPro----- 1003
Db 4037 GGACCGCAGAGGAGCGGCCCATCCCTCCCGCAACCGCCGCCCATCAAGACCTACGA 4096
Qy 1004 -----GlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMethis 1017
Db 4097 GSCCGGAGCCGCATCGGGCTCAAGCTCAAGATCAAGCAGGAGCGGGCTCAGCAAGGT 4156
Qy 1018 ProAla-SerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTy 1037
Db 4157 CGTGACACAACAGGCGCTCGAGCCCGTGACCGGCACCGCCCGCCCGCTACCTCAA 4216
Qy 1037 rValAlaTyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTy 1057
```

```
Db 4217 GGTGGCCGAGCCCCCGCCACCGCCGACCAACCACCG----- 4253
Qy 1057 rGlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaAsgMetMetAl 1077
Db 4253 ----- 4253
Qy 1077 aProProThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHi 1097
Db 4254 -CCGCCACGGCCAGATGAACGACGCGTGACACCCCGCGCTGCGCCCGCCGACGCG 4312
Qy 1097 sGluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerPr 1117
Db 4313 CAAGCCCTGGGCCACCCCGCCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 4361
Qy 1117 oSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAl 1137
Db 4362 -----ACCTACCGGAGAACGCTGGGGGCGCTGGCGC 4393
Qy 1137 aThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGl 1157
Db 4394 G-----CCGAGGGGACGCGCCGACGAGGACG 4423
Qy 1157 nSerGlnHisGlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAl 1177
Db 4424 G-----GGAGGACG-----CCGCGCGCGCTGCGCGCGCGGAGGACGAG 4471
Qy 1177 aAlaGlnAlaLeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLe 1197
Db 4472 CGGCTCATCCGCGAGCTGGCGCGCTGGAGGACGAG-----CTGTACACGCGTATGCT 4525
Qy 1197 u-----AlaProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAs 1215
Db 4526 GAAGGGCCCCCGCCAGAGCCGCGAGCCGCGCGCCCAAGGACACCGGGGACCCGACCTG 4585
Qy 1215 nSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAl 1235
Db 4586 GGAGGCGCGCGGG-----CTGCCCTGCG 4609
Qy 1235 aTyrThrAsnProProHisMetAlaHisValProGlnAla----- 1248
Db 4610 CAAGCGGGCAAGTCCGCGCGCTGCGCGCGAGTGTGACGAGCCAGCTTCTCCAGCGACGCCC 4669
Qy 1249 -----HisValGlnSerGlyMetValProSerHisProThrAl 1261
Db 4670 GCAGGATGACACGCTCAACGAGCACCTGCGAGCGCCATCGACAGCATCTCTGAACCTGCA 4729
Qy 1261 aHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeuAl 1281
Db 4730 GCAGGCGCCC-----GGCCGAGCGCGCGCGCTGCTGATCCCGCCACGCTGCTCGGC 4780
Qy 1281 aGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThrHi 1301
Db 4781 CGGCACCCCGCCATCCCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCTGCGCG 4840
Qy 1301 s 1301
Db 4841 C 4841
RESULT 3
US-11-108-172-1058
; Sequence 1058, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yudi
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
```





Db 6157 ACGGTGACCCCAACCCCAACCCGCGCACACAGACCCCAACCCAGCACCATCACC 6216  
Qy 1042 oGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyr----- 1057  
Db 6217 ACCACCACTA---CGGTGACCCCAACCCCAACCCGCGCACACAGACCCCAACCCAGC 6273  
Qy 1058 -----GlnSerGlnHisProHisValTyrSerPr 1067  
Db 6274 ACACCCATCACCACCACTACCGGTGACCCCAACCCCAACCCGCGCACACAGACC 6333  
Qy 1067 oValIleGlnGlyAsnAlaArgMetAlaProProThr-HisAlaGlnProGlyLeuV 1087  
Db 6334 C-----CAACACGACACCATCACCACACCACTACCGGTGACCCCAACCCCA 6381  
Qy 1087 alSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysP 1107  
Db 6382 ACACCCACCGGCACACAGACCCCAACCCGCGCACACCATCACCACCACTACCGGTGACC 6441  
Qy 1107 roLysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlyS 1127  
Db 6442 CCAACCCCAACACCCGCGCACACAGACCCCAAC-----CAGACACCC 6486  
Qy 1127 erLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHisThr--ProHisPr 1146  
Db 6487 ATCACCACCACTACCGGTGACCCCAACCCCAACCCGCGCACACAGACCCCAACCC 6546  
Qy 1146 oGln---ProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGlyGlySer 1163  
Db 6547 ACGACACCCATCACCACCACTACCGGTGACCCCAACCCCAACCCGCGCACACAG 6606  
Qy 1164 -----HisProAlaProSerProValGlnHisGlnHisGlnAlaGla 1179  
Db 6607 ACCCAACCCACGACACCCATCACCACCACTACCGGTGACCCCAACCCCAACCCACCC 6666  
Qy 1179 nAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaPr 1199  
Db 6667 GGCA---CACAGACCCCAACCCGACAC-----CC 6693  
Qy 1199 oThrProProSerMetThr-----ProAlaSerAsnThrGlnSerPr 1213  
Db 6694 ATCACCACCACTACCGGTGACCCCAACCCCAACCCGCGCACACAGACCCCAACCC 6753  
Qy 1213 oGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHisPro-----SerHi 1231  
Db 6754 ACGACACCCATCACCACCACTACCGGTGACCCCAACCCCAACCCGCGCACACAG 6813  
Qy 1231 sValGlnProAlaTyrThrAsnProHisMet----- 1242  
Db 6814 ACCCAACCCACGACCCCATCACCACCACTACCGGTGACCCCAACCCCAACCCACCC 6873  
Qy 1243 -AlaHisValProGlnAlaHisValGlnSerGly-----MetValProSe 1257  
Db 6874 GGCACACAGACCCCAACCCGCGCACACCATCACCACCACTACCGGTGACCCCAACCCCA 6933  
Qy 1257 rHisProThrAlaHisAlaProMetMetLeuMetThrGlnProProGlyGlyProGl 1277  
Db 6934 ACACCCACCGGCACACAGACCCCAACCCGCGCACACCATCACCACCA----- 6982  
Qy 1277 nAlaAlaLeuAlaGlnSerAlaLeuGln---ProIleProValSerThrAlaHisPh 1296  
Db 6983 -----CTACCGGTGACCCCAACCCCAACCCGCGCACACAG 7020  
Qy 1296 eProTyrMetThrHisProSer 1303  
Db 7021 ACCCAACCCGACACCCATCA 7042

## RESULT 4

US-11-075-185-35/c  
; Sequence 35, Application US/11075185  
; Publication No. US20050266434A1  
; GENERAL INFORMATION:  
; APPLICANT: REEVES, CHRISTOPHER D  
; APPLICANT: JULIEN, BRYAN

; APPLICANT: REID, RALPH  
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS  
; FILE REFERENCE: 010099.03  
; CURRENT APPLICATION NUMBER: US/11/075,185  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: US 60/551,103  
; PRIOR FILING DATE: 2004-03-08  
; PRIOR APPLICATION NUMBER: US 60/568,290  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 35  
; LENGTH: 10968  
; TYPE: DNA  
; ORGANISM: Sorangium cellulosum  
US-11-075-185-35

## Alignment Scores:

Pred. No.: 5 08e-08 Length: 10968  
Score: 437.00 Matches: 353  
Percent Similarity: 32.24% Conservative: 149  
Best Local Similarity: 22.67% Mismatches: 592  
Query Match: 6.28% Indels: 470  
DB: 7 Gaps: 66

US-10-802-228-2 (1-1312) x US-11-075-185-35 (1-10968)

Qy 18 ArgArgPheAlaAlaAlaAlaArgTrrProGlyTrrArgSerLeuGlnArgProAlaArgArg 37  
Db 7135 CGTGGAGCGCGCGCGGAGAGCGGGTGGGTGGCGTACAGGCGCGGCCCATCGCGCGC 7076  
Qy 38 SerGlyArgGlyGlyGlyAlaAlaProGlyTrrProSerAlaAlaPro----- 55  
Db 7075 GCTGGTGCCTCGCGCGGTGAAGAGCACCAGCAGCTTG-CCGCGCGCGCTCCCGCGCGC 7017  
Qy 56 -----ProProGlyProGlyProProPro-----SerArgGlnSerSerPro 70  
Db 7016 GTCACGCGCGCGCGCTGGCGCGCGCGCGCTGAGCGCAACTCGGCAAGCGCGCGCGC 6957  
Qy 71 ProSerAlaSerAspCysPheGlySerAsnGlyGlyGlyAlaPheArg----- 88  
Db 6956 AGCTCTCTCGGTGCTGCTCGCGCGCGAGCGCGCGAGCGCA--GCCGCGTGGCGAGGTGC 6899  
Qy 89 -----ProGlySerArgArgLeuLeuGlyLeuGlyPro----- 100  
Db 6898 TGGCGCTGCTGGCGCAGGCTCGCGCGCAGCTCGAGCAGCGCTGCTCGCGGTGGCGCGCA 6839  
Qy 101 ---ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAla 119  
Db 6838 GGTGCTCGG-----CGAGCGCGCGCGCTGGCGC-----GCA 6806  
Qy 120 AlaProThrArgAlaSerProLeuGlyAlaArgAlaSer-----ProPro 134  
Db 6805 GCGCGCGCTGCTCGCGACCCGACAGAGAGCAGCGCGCAGCGCACCGCGCGCGCGC 6746  
Qy 135 ArgSerGlyValSerLeuAlaArgProAlaProGly----- 146  
Db 6745 CTTCCGCGCGCGTCTGCGCGCGCGCGCGCTCTCTCGACGATGACGTGGCGGTGGTGC 6686  
Qy 147 -----Cys-----ProArgProAla-----Cys-GluProValTy 156  
Db 6685 CGTGTGATGCCAAGGAGCAGACCGCGCGCGCGCGCGTCCGTCCGCGCGCGCACCGCA 6626  
Qy 156 rGly-----ProLeuThrMetSerLeuLysProGlnGlnGlnGln----- 170  
Db 6625 CGGGCTCGTTCAGGAGGTTCGTGCGCCCTGCGACCACTCGATGTCGGCGCATCGGGCT 6566  
Qy 171 -----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPr 188  
Db 6565 CGGCATACAGCGTCTTGGCAGCTCTCGTGTCTGATCGCATCGCGCAGCAGCATTCATCAGCC 6506  
Qy 188 oProProAlaAlaAlaAsnValArgLysPro----- 198



Db 6505 CCACAGCGCCCGCGCGCTCGTGTCGCCGAGGTTGGACTTGATCGAGCGACGCCAGA 6446  
Qy 199 -----GlyGlySerGly-----Leu-LeuA 205  
Db 6445 GCGCGCGCTCGCGGTTGTGCGGGTGCCTAGTTCGCCAGCAGCGCTCGCGCTCGATGG 6386  
Qy 205 laSerProAlaAlaLaProSerProSerSerSerValSerSerSerAlaThrA 225  
Db 6385 GATCGCCGAGGCGCGTGCAGTGCCTGCGCTCGACGCGCTCGATGTCGCGCGGGTCA 6326  
Qy 225 laProSerSerValAlaLa----- 231  
Db 6325 GCGCGCGCGCGCGAGGCGCTCGCGGATGACCTCTGTGGCGCGCGGCGTTGGGGCGG 6266  
Qy 232 -----AlaThr-SerGlyGlyArgProGlyLeu 241  
Db 6265 TGAGCCCTTGCTGCGCGCTCTGTTGACCGCGGAGCGCGGAGGACGCCACACGC 6206  
Qy 242 GlyArgGlyArgAsnSerAsn-----LysGlyLeuProGlnSerThr 255  
Db 6205 GGTCCGCTCGCGCGCGCTCGACAGCGCTTGAGCACACAGAGC-CCGACGCCCTCC 6147  
Qy 256 lleSerPheAspGlylleTyrAlaAsnMetArgMetValHisIleLeuThrSerValVal 275  
Db 6146 GCGGGCGCATCGGTTCCGCCCGCCGGAAGGCTTGCACCGTGCCTCGGGGCCATG 6087  
Qy 276 GlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThr 295  
Db 6087 ----- 6087  
Qy 296 TyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSer 315  
Db 6086 -----CCGCGTGTGCGGTGAA-CTCATGAAGGCCATGGCGTGGCCATCACCGTGGC 6034  
Qy 316 SerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheVal 335  
Db 6033 GCGCGCGGAGGCGAGTGCACCTCGCGCGCG----- 5998  
Qy 336 ValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSer 355  
Db 5997 -----GAGCGCATGTCACGCGAGGTGGAGGACACAGCGA 5962  
Qy 356 AlaIleSerAlaLysValAsnGlyGlu-----HisLysGluLysAsp 369  
Db 5961 CGACGACGACGCGTGTCTGACGCGTACCGCGCGCGCGCGAGCCCGCTAGCGCAT 5902  
Qy 370 LeuGluProTyrAspAlaGlyGluLeu-----ThrAlaAsnGluLeuGlu 385  
Db 5901 GCGCGCGGACGCGCTGGGAAGCTGCGGTGGCGAGCTGAGCGTCAAGACCTCGAG 5842  
Qy 386 AlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMetPheArgTyrAsnGlu 405  
Db 5841 GTGCTCGAGACCGCCCGCC----- 5821  
Qy 406 GluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSerTyrThrValProLeu 425  
Db 5820 -----GTAGTCGCTGTACACCGCCGACGAGAGCGCGGTGGCGCTCGCCTC 5773  
Qy 426 GluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAlaA-----AsnGlnLeuAla 444  
Db 5772 GAGCGA-----ATGGGGCGGAGGCGCGCGCTCGAGCGCGTGGCGCTCGCCTC 5731  
Qy 445 GluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspArg 464  
Db 5730 GCGCGACTCGACGACGCGCTGTGCGGGTCCATGCGCTCCGCTTCGCGCGGCTGAT 5671  
Qy 465 SerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArgGluGly----- 482  
Db 5670 GCCGAAGAAGTTCGATCGAAGCGGT-CAGCGCTCGTAGGAAGCGCGCCAGGTTGGTGA 5612  
Qy 483 -----HisIleAsnThr 487  
Db 5611 CCGACTTCCCGGGGCGTCCGGTTCGGGGTGTATACAGCCCGCCACATCCCGAGCGCGCC 5552

Qy 488 -----ArgGluAsnLysTyrIleProPro 495  
Db 5551 CCTCGGGGAACGGCCCGATGCGTCCCGGCCCTCCGAGAGCAGCTGCCACACGCGCGCG 5492  
Qy 496 GlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMet 515  
Db 5491 GGGTATCACGCGCGCGCGGAGCGGCATCGACACGATCGCATCGCTCGCTCGCT 5432  
Qy 516 GlyGlnPro-----GlySerGlySerMetProSerArgSerThr 528  
Db 5431 GCGCTCCCGCGCGCGCGGCGCTGTGCGCGCTCGCCCGCGCGCGAGCTCCGAGA 5372  
Qy 529 SerHisThrSerAspPheAsnProAsnSer-----Gly 539  
Db 5371 ACGCAGCTGAAGCAGCAGATCCCGCATCGCGCGCGCTCGGATGTTGTAACGCGAGG 5312  
Qy 540 Ser-----AspGlnArgValValAsnGlyGlyValProTyrProSerPro 554  
Db 5311 TCGCGCGCAACGCGCTCTCGCGCGCGCGCGAGCGGTTCCTCAGCTCGACGCGCTGA 5252  
Qy 555 -----CysProSerProSerArgProProProSerArgTyrGlnSerGlyProAsn 571  
Db 5251 GCGAGGTGAGCCCGCTCTTCAGCGCTTGTCCG-----CGGGGACCTCGG 5204  
Qy 572 SerLeuProProArgAlaAlaThrProThrArgProProSerArgProProSerArgPro 591  
Db 5203 CCGACCTGCCAGCTGACGACGCGCG-----CGGCTCCGCGCGCACCGCTCGAGCA 5150  
Qy 592 SerArgProProSerHisProSerAlaHisGlySerProAlaProValSerThrMetPro 611  
Db 5149 GCGCGCGCGCGCTCGGCTCGGCA-----GGCGCG----- 5117  
Qy 612 LysArgMetSerSerGluGlyProProArgMetSerProLysAlaGlnArgHisProArg 631  
Db 5116 ---AGAGCGCGCGCGAGCGCGCTCGGTCGCGCGCGCGCGCGAGCGCGCGCG 5060  
Qy 632 AsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHis 651  
Db 5060 ----- 5060  
Qy 652 AsnProProSerGluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThr 671  
Db 5059 -----GCG 5024  
Qy 672 TrpSerSerValSerGlyValProArgLeuSerProLysThrHisArgProArgSer 691  
Db 5023 -----GCGCGCGCACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4997  
Qy 692 ProArgGlnAsnSer---IleGlyAsnThrProSerGlyProValLeuAlaSerProGln 710  
Db 4996 CCGCAGCTGAGCG 4943  
Qy 711 AlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAlaSerPro---Thr 729  
Db 4942 GCGCGCGAGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4883  
Qy 730 ProAla-----SerProAlaSerAsnArgAlaValThrProSerSerGluAlaLys 746  
Db 4882 CTGCGCTCTGAGCGGTGCCAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4823  
Qy 747 AspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnLys 766  
Db 4822 CTGCTCCCAAGAGCCCGAGGAGGCTCATCGCGCGCGCGCGCGCGCGCGCGCGCG 4763  
Qy 767 Pro-----AsnGluThrSerPro-SerPheSe 775  
Db 4762 CCGGAGCGCGCTCCAGGAACGCTTGGCGCGCGGTAGTTCCGCTCGCGCGCGCTGCCA 4703  
Qy 775 rLysAla-----GluAsnLysGlyIleSerProValSerGluHisAr 790  
Db 4702 GCGTCCGCGCGCGCGAGAGAACAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 4643

QY 790 glysGlnIleAspAspLeuLysLysPhelYsAsnAspPheArgLeuGlnProSerSerTh 810  
DB |||:|||||  
4642 CGGCCAGGTGCCAGCGCC-----GTCGATCTTCGGCCCGCAGCACGC 4601  
QY 810 rSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLe 830  
DB |||:|||||  
4600 GCGGAGTGTCTC-----GGCGGAGAGT 4577  
QY 830 uLeYsAspLysIleGluProSerAlaLysAspSerPheIleGluAsn-SerSerSera 850  
DB |||:|||||  
4576 TCGTCACGATGAGTGTCCAGCACGCCGC-----CATGTGCACGACCGCGTGGGG 4523  
QY 850 snCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSera 870  
DB |||:|||||  
4522 GGTGTGCACCGCGGATCGCGCGATACCGCCCGCAGCGCGCGCGTGGCG----- 4470  
QY 870 snThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSer-Pro 889  
DB |||:|||||  
4469 --ACGTGCACCGCGCGAGTGCACCGTTCGGCGCGCGCGCGCGCGTGTCTCACCA 4412  
QY 890 AlaCysLysGlnGluLysAspAspLysGluGlu-----LysLysAspAlaAlaGlu 906  
DB |||:|||||  
4411 GCGCGCGCGTCCGCGCATCCATCCGCGCGCGAGTGAGCACGAGTGCAGCACGC 4352  
QY 907 Gln-----ValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
DB |||:|||||  
4351 CGTGCAGCGCGACAGGTGCGCGCGACCTCGCGCGCGAGTCTCCCGGTGCGCGCGTGA 4292  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
DB |||:|||||  
4291 TCGACCGTTCGCGCGGATCCAGCGCGTTCGTCGCGCGCGTTCGCGCGCGCGCGCG 4232  
QY 941 Ala-----GlnProSerProSerMetValGlyHisGlnGlnPro 953  
DB |||:|||||  
4231 GCACCTCGCGCAGCGTGGCGCGCGCGCGCGCGCGCGCGCGAGTTCGGTCTCT 4172  
QY 954 ThrProValThrThrGlnProValCysPheAlaProAsnMetMetThrProValProVal 973  
DB |||:|||||  
4171 CGCGCG-----CGGCCA 4160  
QY 974 SerProGlyValGlnProLeuThrProLeuProMetThrPro-----MetPro 989  
DB |||:|||||  
4159 CGCGCGCGCGCGCGCGCGTTCGTCAGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCG 4100  
QY 990 ValAsnGlnAlaLysThrThrArgAlaValProAsnMetProGlnGlnArgGlnAspGln 1009  
DB |||:|||||  
4099 CGCGCGCGGTTTC-TCTGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4059  
QY 1010 HisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThr 1029  
DB |||:|||||  
4058 ---CACAGCGCGTGGTCAACCGCGCGCGCGCGCGTTCGTCGCGCGCGCGCGCGCGCG 4002  
QY 1030 ProProAlaThrSerThrGlnThrVal-----1038  
DB |||:|||||  
4001 CGCGCGTCAACCAACCACTCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942  
QY 1039 -----AlaThrSerProGlnGlnPheProAsn 1047  
DB |||:|||||  
3941 ACCAGCGACAGGACCGTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3882  
QY 1048 GlnProLeuValGlnHisValProHisThrGln---SerGlnHisProHisValThrSer 1066  
DB |||:|||||  
3881 TCGCGCGCGCGCGTTCGATCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3822  
QY 1067 ProValleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGln-----1083  
DB |||:|||||  
3821 GCGATGAGTCCG 3762  
QY 1084 ---ProGlyLeuValSerSer-----1090  
DB |||:|||||  
3761 CGCGCGAGCAGTGTCTCTCG 3702  
QY 1091 ---AlaThrGlnThr-GlyAlaHisGluGlnThrHisAlaMetThrAlaCysProLysLe 1109

DB |||:|||||  
3701 CACGTGACCGGTACAGGT---CAGCTCGTCCGACCCCGCGCGCGTTCGCGCGCGTGC 3645  
QY 1109 uProThrAsnLysGluThrSerProSerPheThrPheAlaIleSerThrGlySerLeuAl 1129  
DB |||:|||||  
3644 TCCG---CCGTGCG 3588  
QY 1129 aGlnGlnThrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSe 1149  
DB |||:|||||  
3587 GCGTGGCGCGTTCGATCGACCGCGCGAGTGGTCCCACTCGTGGCGTCTCGACCTCC 3528  
QY 1149 rAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaProSerPr 1169  
DB |||:|||||  
3527 AGCGCGCGCGCGCGCGCGCGTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3480  
QY 1169 oValGlnHisHisGlnHis-----GlnAlaAlaGlnAlaLeuHisLeuAlaSerPr 1186  
DB |||:|||||  
3479 GACCACCGGAAGGATCGGATGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3420  
QY 1186 oGlnGlnGlnSerAlaIleThrHis-----AlaGlyLeuAlaProThrPr 1201  
DB |||:|||||  
3419 CGCAGGAGCGCGTGGAGCGCGCGCATCGACAAAGCGCGGATGCAAGCGCGAGTCTCGCGC 3360  
QY 1201 oProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAla-----1219  
DB |||:|||||  
3359 CGC-TCCGAGACCCCTCGGGAACCGCCAGGTGCGGTAGATCGACGCGCGCGCGCGCGCG 3301  
QY 1220 -AlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaThr-----1237  
DB |||:|||||  
3300 CGCGCGCGTACG 3241  
QY 1238 -----AsnProHisMetAlaHisValProGlnAlaHis-----1249  
DB |||:|||||  
3240 CGTGCAGTGCAGCGGTCTCG 3181  
QY 1250 -ValGlnSerGlyMetValProSerHisProThrAla-----His-----1262  
DB |||:|||||  
3180 GCG 3121  
QY 1263 -AlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeuAlaG1 1282  
DB |||:|||||  
3120 CGCGCGTTCGCGCGCGTGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3061  
QY 1282 nSerAlaLeuGlnProIleProValSerThrThrAla 1294  
DB |||:|||||  
3060 CACC-----GTCACCTGCAGCGTCAACCGCG 3036

## RESULT 5

US-11-075-185-1/c  
; Sequence 1, Application US/11075185  
; Publication No. US20050266434A1  
; GENERAL INFORMATION:  
; APPLICANT: REEVES, CHRISTOPHER D  
; APPLICANT: REID, RALPH  
; APPLICANT: JULIEN, BRYAN  
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS  
; FILE REFERENCE: 010099.03  
; CURRENT APPLICATION NUMBER: US/11/075,185  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: US 60/551,103  
; PRIOR FILING DATE: 2004-03-08  
; PRIOR APPLICATION NUMBER: US 60/568,290  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 78869  
; TYPE: DNA  
; ORGANISM: Sorangium cellulosum  
US-11-075-185-1

Alignment Scores:

Pred. No.: 2.74e-07 Length: 78869

Score:	437.00	Matches:	353
Percent Similarity:	32.24%	Conservative:	149
Best Local Similarity:	22.67%	Mismatches:	592
Query Match:	6.28%	Indels:	470
DB:	7	Gaps:	66

US-10-802-228-2 (1-1312) x US-11-075-185-1 (1-78869)

Qy	18	ArgArgPheAlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArg	37
Db	22402	CGTCGAGCGCGCGCGAAGACGGGTGGTGGCTACAGGGCGGCCCCATGCGCGGC	22343
Qy	38	SerGlyArgGlyGlyGlyAlaAlaProGlyProTyProSerAlaAlaPro-----	55
Db	22342	GCTGGCTGCTCCCTGGCGGTGAAGAGCACCACCTTG-CCGGGGGGTTCGCCCGCGCG	22284
Qy	56	-----ProProGlyProGlyProProPro-----SerArgGlnSerSerPro	70
Db	22283	GTACGGCGCGCGCTGGCGCGCGCGCTGAGGGAACTCGGCAAGCGGGCGGCC	22234
Qy	71	ProSerAlaSerAspCysPheGlySerAsnGlyGlyGlyAlaPheArg-----	88
Db	22223	AGCTCTCGTGGTGGTCCGCGCGCGAGCGGCGCA--GCCGCGTGGCGAGGTGG	22166
Qy	89	-----ProGlySerArgArgLeuLeuGlyLeuGlyGlyPro-----	100
Db	22165	TGCGGGTCTGGTCCAGGCTCGCGCGCGACGTCGAGCAGCGCGTGGTCCGGTGGCGCGCA	22106
Qy	101	---ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProAla	119
Db	22105	GGTGTCTGG-----CGAGCGCGCGCGCTCGGGCGC-----GCA	22073
Qy	120	AlaProThrArgAlaSerProLeuGlyAlaArgAlaSer-----ProPro	134
Db	22072	GGCGCGGTCTGCGGACCCGACAGGAGCAGCGGCGAGCGTCAACGGCAGCGCGCGCG	22013
Qy	135	ArgSerGlyValSerLeuAlaArgProAlaProGly-----	146
Db	22012	CCTCGCGCGCGCTCTCGCGCGCGCGCGCTCTCGACGATGACGTGGGTGGTGC	21953
Qy	147	-----Cys-----ProArgProAla-----Cys-GluProValTy	156
Db	21952	CGCTGATCCGGAAGACGACGACCGCGCGCGCGCGCGCGTCCGTGGCGGCCACGGCA	21893
Qy	156	rgly-----ProLeuThrMetSerLeuLysProGlnGlnGlnGln-----	170
Db	21892	CGGGCTCGTTACGAGGTGTGTCCTCGACACGTCGATGTGGCGGATCGGGCT	21833
Qy	171	-----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPr	188
Db	21832	CGGCATACAGCTCTTGGCAGCTCTGGCTGTGTCATCGCACGACGACGCTTCATCAGCC	21773
Qy	188	oProProAlaAlaAlaAsnValArgLysPro-----	198
Db	21772	CCACGACGCGCGCGCGCTGCGTGGTCCGAGTTGGACTGTGATCGAGCGGACCCAGA	21713
Qy	199	-----GlyGlySerGly-----Leu-LeuA	205
Db	21712	GGCGCGCTCGCGGTGGCGGTGGCTAGTGCACGACGCGCTGCGCTCGATGG	21653
Qy	205	laSerProAlaAlaAlaProSerProSerSerSerSerSerSerSerSerSerSerSer	225
Db	21652	GATCGCCGAGGCGGTGCGAGTGGCTGCGCTCGACGCGCTCGATGTCGCGCGGGTCA	21593
Qy	225	laProSerSerValValala-----	231
Db	21592	GCCCCGCGCGCGCGCGCTGGCGGATGACGTCCTGTGGCGCGCGCGTGGGGCGG	21533
Qy	232	-----AlaThr-SerGlyGlyArgProGlyLeu	241
Db	21532	TGAGCCCTTGGCTGGCGCGTCTGTTGACCGCGGAGCGCGGAGGACGCGCACACGC	21473
Qy	242	GlyArgGlyArgAsnSerAsn-----LysGlyLeuProGlnSerThr	255

Db	21472	GGTCGCGCTCGCGCGCGCTCCGACAGCCGCTTGAGCACCAGGAGC-CCGACAGCCCTCC	21414
Qy	256	IleSerPheAspGlyIleTyAlaAsnMetArgMetValHisIleLeuThrSerValVal	275
Db	21413	CGGGGCGCATGCTCGCTCGCGCGCGCCCGAAGGCTTGACCGTGGCTCCGGGGCCATG	21354
Qy	276	GlySerLysCysGluValGlnValLysAsnGlyGlyIleTyArgGluGlyValPheLysThr	295
Db	21354	-----	21354
Qy	296	TyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSer	315
Db	21353	-----CCGCGCTGTGGCTGAA-CTCGATGAAGGCCATGGGCGTGGCCATCACGTTGC	21301
Qy	316	SerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheVal	335
Db	21300	GCGCGCGGAGGCGCATGTCGCTCGCGCGCG-----	21265
Qy	336	ValValGlnPheLysAspMetAspSerSerTyAlaLysArgAspAlaPheThrAspSer	355
Db	21264	-----GAGCGACATGCACCGAGGTGGAGGACACACGCGA	21229
Qy	356	AlaIleSerAlaLysValAsnGlyGlu-----HisLysGluLysAsp	369
Db	21228	CGACGAGCAGCGCTGTGACGCGTGCACCGCGCGCGCGGCGGAGCCCGACGCTGTAGCCGAT	21169
Qy	370	LeuGluProTrpAspAlaGlyGluLeu-----ThrAlaAsnGluLeuGlu	385
Db	21168	GCGCGCGGAGCGCTGGGAAAGCTCCGCGTGGCGACGTAGCGTGAAGACCTCGAG	21109
Qy	386	AlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyArgAsnGlu	405
Db	21108	GTGCTCGAGACGCCCGCC-----	21088
Qy	406	GluAsnTyGlyValValSerThrTyArgSerSerLeuSerSerTyThrValProLeu	425
Db	21087	-----GTAGTCGCTGTACACGAGCCGAGGAGCGCGCGTGGCGCTCGCCTC	21040
Qy	426	GluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla--AsnGlnLeuAla	444
Db	21039	GAGCGA-----ATGGGCGCGAGGCGCGCGCTCGAGCGCTCCCA	20998
Qy	445	GluGluIleGluSerSerAlaGlnTyLysAlaArgValAlaLeuGluAsnAspArg	464
Db	20997	GCGCACTCGAGCAGCAGCGCTGCTGGGTCCATCGCTCCGCTTCGCGCGGCTGAT	20938
Qy	465	SerGluGluGluLysTyThrAlaValGlnArgAsnSerSerGluArgGly---	482
Db	20937	GCCGAAGAAGTGGATCGAAGCGGT-CAGCGTCTGTAGAGGAAGCCGCCAGTTGGTCA	20879
Qy	483	-----HisSerIleAsnThr	487
Db	20878	CCGACTTCCCGCGCGCTGTCGGGTGCGTACAGCCCGCCACATCCCGAGCGCGCC	20819
Qy	488	-----ArgGluAsnLysTyIlePro	495
Db	20818	CCTCGGGGAAAGCGCGGATGGGTCCCGGCCCTCGAGAGCAGCTGCCACAGCGCGCG	20759
Qy	496	GlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMet	515
Db	20758	GGGTATCAGCGCGCGCGGAGCGGCACGCGCATCGACAGCTCGCATCGCTCGCTCGT	20699
Qy	516	GlyGlnPro-----GlySerGlySerMetProSerArgSerThr	528
Db	20698	GCGTCTCCCGCGCGCGCTGTCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG	20639
Qy	529	SerHisThrSerAspPheAsnProAsnSer-----Gly	539
Db	20638	ACGACGCTGAAGCAGCAGATCCGCGATCGCGCGCGCGCTCGATGGTTCGAACGCGAGG	20579
Qy	540	Ser-----AspGlnArgValValAsnGlyGlyValProTrpProSerPro---	554

Db 20578 TCAGCGGAAACGGCTCTCGGCGCGCGCGCGGAGGGCGGTTCCTCAGCTCGACGGCGGTGA 20519  
QY 555 -----CysProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsn 571  
Db 20518 GCAGGTGAGCCAGCTCTTCAGCGGCTGTTCGCG-----CGGGACCTCGCG 20471  
QY 572 SerLeuProProArgAlaAlaThrProThrArgProProSerArgProProSerArgPro 591  
Db 20470 CCGGACCTGCCAGCTCGACACGCGCG-----CCGCTTCGCGCGCACACGAGCTCGAGCA 20417  
QY 592 SerArgProProSerHisProSerAlaHisGlySerProAlaProValSerThrMetPro 611  
Db 20416 GCAGCGCGCGCGCTCGGCTCGGCA-----CGCGCG----- 20384  
QY 612 LysArgMetSerSerGluGlyProProArgMetSerProLysAlaGlnArgHisProArg 631  
Db 20383 ---AGAGCGCGCGCGCGCGCGCTCGGTTCGCGCGCGCGCGCGCGCGCGCGCG 20327  
QY 632 AsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHis 651  
Db 20327 ----- 20327  
QY 652 AsnProProSerGluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThr 671  
Db 20326 -----GCCG 20291  
QY 672 TrpSerSerValValSerGlyValProArgLeuSerProLysThrHisArgProArgSer 691  
Db 20290 -----GCCG 20264  
QY 692 ProArgGlnAsnSer---IleGlyAsnThrProSerGlyProValLeuAlaSerProGln 710  
Db 20263 CCGCAGCTCGAGCG 20210  
QY 711 AlaGlyIleThrProThrGluAlaValAlaMetProIleProAlaAlaSerPro---Thr 729  
Db 20209 GCGCGCGAGCG 20150  
QY 730 ProAla-----SerProAlaSerAsnArgAlaValThrProSerSerGluAlaLys 746  
Db 20149 CTGCGCGTCTGAGCGGTGCGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 20090  
QY 747 AspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAlaLys 766  
Db 20089 CTGCTCCAGAACCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 20030  
QY 767 Pro-----AsnGluThrSerPro-SerPheSe 775  
Db 20029 CCGCGAGCGGTCCAGGAACGGTTCGCGCGCGGTAGTTCGCGCGCGCGCGCGCGCGCG 19970  
QY 775 rLysAla-----GluAsnLysGlyIleSerProValValSerGluHisAr 790  
Db 19969 CGGTGCG 19910  
QY 790 gLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerTh 810  
Db 19909 CGGCGAGGTGCAGCGCGCG-----GTGATCTTCGCGCGCGCGCGCGCGCGCGCG 19868  
QY 810 rSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLe 830  
Db 19867 GCGCGAGCTGCTC-----GCCGCGAGAGCT 19844  
QY 830 uLleLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsn-SerSerSerA 850  
Db 19843 TCGTACAGATGAGCTGCTCGACGAGCGCGCG-----CATGTGACACACCGCGCGTGA 19790  
QY 850 snCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerA 870  
Db 19789 GGTGTGCGCGCGCGATCGCGCGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19737  
QY 870 snThrGluHisArgGlyProGluValThrSerGlnGlyValGlnThrSerSer-Pro 889  
Db 19736 --ACGTGCGCAGCG 19679

QY 890 AlaCysLysGlnGluLysAspAspLysGluGlu-----LysLysAspAlaAlaGlu 906  
Db 19678 GCG 19619  
QY 907 Gln-----ValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920  
Db 19618 CGTGCG 19559  
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940  
Db 19558 TCAGCACCGCTCCCGCGCGATCCAGCGCGCTCGGTTCGCGCGCGCGCGCGCGCGCG 19499  
QY 941 Ala-----GlnProSerProSerMetValGlyHisGlnGlnPro 953  
Db 19498 GCACCTCG 19439  
QY 954 ThrProValThrThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal 973  
Db 19438 CGCGCG-----CGGCCA 19427  
QY 974 SerProGlyValGlnProLeuTyrProIleProMetThrPro-----MetPro 989  
Db 19426 CCG 19367  
QY 990 ValAsnGlnAlaLysThrTyrArgAlaValProAsnMetProGlnGlnArgGlnAspGln 1009  
Db 19366 CCGCGCGCGCGGTTC-TCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19326  
QY 1010 HisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThr 1029  
Db 19325 ---CACAGCGCGCGGTGGCTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19269  
QY 1030 ProProAlaTyr-SerThrGlnTyrVal----- 1038  
Db 19268 CCGCGCGTCACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19209  
QY 1039 -----AlaTyrSerProGlnGlnPheProAsn 1047  
Db 19208 ACCAGCGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19149  
QY 1048 GlnProLeuValGlnHisValProHisTyrGln---SerGlnHisProHisValTyrSer 1066  
Db 19148 TCAGCG 19089  
QY 1067 ProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGln----- 1083  
Db 19088 GCGATGAGCTCCG 19029  
QY 1084 ---ProGlyLeuValSerSerSer----- 1090  
Db 19028 CCGCGCGACGAGTGTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18969  
QY 1091 ---AlaThrGlnTyr-GlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLysLe 1109  
Db 19968 CAGTGACCGGTACAGGT---CACGCTGCTCGCACCGCGCGCGCGCGCGCGCGCG 18912  
QY 1109 uProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAl 1129  
Db 18911 TCCG---CCGTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18855  
QY 1129 aGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSe 1149  
Db 18854 GCGTCGCGCGAGCTCGATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18795  
QY 1149 rAlaThrProThrGlyGlnGlnGlnSerGlnHisGlyGlySerHisProAlaProSerPr 1169  
Db 18794 ACG 18747  
QY 1169 oValGlnHisHisGlnHis-----GlnAlaAlaGlnAlaLeuHisLeuAlaSerPr 1186  
Db 18746 GACCGCGAGGCGATCGGGATGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 18687

QY 1186 cGlnGlnSerAlaIleTyrHis-----AlaGlyLeuAlaProThrPr 1201  
Db 18686 CGCAGAGGGCTGAGCGCCATCGAGCAAGCGCGATGCAAGCCGAAGTCTCTGCGCC 18627  
QY 1201 oProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAla----- 1219  
Db 18626 CGC-TCCGAGACCCCTCGGGCAAGCCAGGTGCGGTAGATCGACCGCCCGCCACCGGCA 18568  
QY 1220 -AlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThr----- 1237  
Db 18567 CGCCCGCTGACGCCCCCGAAGCGCGGCGGACCCACCGCTCGAGCGCTCGACCCACCG 18508  
QY 1238 -----AsnProProHisMetAlaHisValProGlnAlaHis----- 1249  
Db 18507 CGTGACGTGACCGGCTCGGCCCGCGCGCGCGCCACCGCGCCCGCCACCGCGCTCGT 18448  
QY 1250 -ValGlnSerGlyMetValProSerHisProThrAla-----His----- 1262  
Db 18447 GCGCGCGGGCGGTGCGCCCGAGCAGCGCGCGCGGTGCGCGCTCGCGCGCGCTCGT 18388  
QY 1263 -AlaProMetMetLeuMetThrThrGlnProProGlyProGlnAlaAlaLeuAlaG1 1282  
Db 18387 CGCGCGCTCGGCGCGCTGTGAGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCTC 18328  
QY 1282 nSerAlaLeuGlnProIleProValSerThrAla 1294  
Db 18327 CACC-----GTCACCTGACGCGTCAACCGCC 18303

## RESULT 6

US-11-056-470-1  
; Sequence 1, Application US/11056470  
; Publication NO. US20050261263A1  
; GENERAL INFORMATION:  
; APPLICANT: SANTI, Daniel V.  
; APPLICANT: MYLES, David C.  
; APPLICANT: TIAN, Zong-Qiang  
; APPLICANT: HUTCHINSON, C. Richard  
; APPLICANT: JOHNSON, Robert G., Jr.  
; APPLICANT: ZHOU, Yi-Qing  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: BENZOQUINONE ANSAMYCINS  
; FILE REFERENCE: 300622007510  
; CURRENT APPLICATION NUMBER: US/11/056,470  
; CURRENT FILING DATE: 2005-02-11  
; PRIOR APPLICATION NUMBER: US 10/212,962  
; PRIOR FILING DATE: 2002-08-05  
; PRIOR APPLICATION NUMBER: US 60/310,779  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/389,255  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: US 60/393,929  
; PRIOR FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/395,275  
; PRIOR FILING DATE: 2002-07-12  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; TYPE: DNA  
; ORGANISM: Streptomyces geldanus  
US-11-056-470-1

## Alignment Scores:

Pred. No.:	3,74e-08	Length:	6210
Score:	435.00	Matches:	337
Percent Similarity:	32.98%	Conservative:	158
Best Local Similarity:	22.45%	Mismatches:	491
Query Match:	6.25%	Indels:	519
DB:	7	Gaps:	67

US-10-802-228-2 (1-1312) x US-11-056-470-1 (1-6210)

QY 9 ArgSerProAlaValAlaThrGluSerArgPheAlaAlaAlaArgTrpProGly--- 27

Db 101 CGATCGCGCGGT-----TCCGGAGGACCGGGGCTGGACCTGCGCGGGCTCT 148  
QY 27 ----- 27  
Db 149 TCGACTCGACCCGAGCGCCGTGGCAAGTCTATGTGCGCGAGGGCGGTTTCTTCACCG 208  
QY 27 ----- 27  
Db 209 ACGGGGCGGATTGACGCGGCATTTCTTCGGCATCTCGCCCGTGAAGCGTGGCGATGG 268  
QY 28 -----TrpArgSerLeuGlnArgProAla----- 35  
Db 269 ACCGCGACAGCGGTGCTGCTGGAGACCGCGTGGAGACCTTCAGAAATCCCGAATCG 328  
QY 35 ----- 35  
Db 329 ACCCGGTTCTCTGACGGCACCGACGTCGGTGTGTTTCAGCGGAGTGATGTACCACGATT 388  
QY 36 -----ArgArgSerGlyArgGly----- 41  
Db 389 ACGGGCGCGACCGCGGACGCGCGGAGGCGCTGGAGGGCATCTCGCGCTGGCGACCG 448  
QY 42 GlyGlyGlyAlaAlaProGlyProTyrProSer-----AlaAlaProProProProGly 59  
Db 449 CGGGAGCGTCTCTCGGGCGGTGCGCTACGGCGTGGGCTGACCGGCGCGCGGTGA 508  
QY 60 ProGlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySer 79  
Db 509 CCGTGGACACCGCTGCTCGT-----CCTCCTT--GGTAGCGTGACCTGGCGGT 557  
QY 80 AsnGlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGly 99  
Db 558 TCAGCGGTGGCGACCGGCGGAATGCTCGTGGCTGCGCG-----GGGTGTGCGGT 611  
QY 100 ProProArgProPheVal-----ValLeuLeuPro--LeuAlaSerProGlyAlaP 117  
Db 612 GATGAGCAGCGCGACGCTGCTTCATCGAGTTCTCCCGCAGCGTGGCTCGCCCGCGCG 671  
QY 117 roProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerG 137  
Db 672 CCGTGCAGTCTTCGGAGGGCGCGAG-----GCACCACTGGTCCGAGGGTGTGCG 728  
QY 137 lyValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrG 157  
Db 729 GTTGGTGTTC-----TGGAGCGGTCTCGATGCCCGCC----- 763  
QY 157 lyProLeuThrMetSerLeuLys--ProGlnGlnGlnGlnGlnGlnGlnGlnGln 176  
Db 764 -----GCAATGGGATGAGGTGCTCGCGCTGCTCGTGGCACCGCGCGTGAACCGGA 815  
QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196  
Db 816 CGGCGCCAGCAACCGCTGACCGCGCCACCGCGCGTCCCGAGAACGGGTGATCCGCGCA 875  
QY 197 LysProGly-GlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSe 216  
Db 876 GCGCTGGCGAACCGCGGCTGACGGTGGCGCATGTGGACCGGTGAGAGCGCCACCGCAC 935  
QY 216 rSerValSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyG1 236  
Db 936 GGGCACGAGTCTCGCGCACCGCTGAGAGCGCCAGGCGCATCTCTG---GCCACTACCGGCA 992  
QY 236 yGlyArgPro-----GlyLeuGlyArgGlyArgAsnSerAsnLysGlyLe 251  
Db 993 GGAGCGCGCGGAGGATCAGCGCTGTGCTGGTGGG---TCGTTGAAGTGAACATCGCGCA 1049  
QY 251 uProGlnSerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLe 271  
Db 1050 TGCGCAGCGCGCGCGCGCGCGGTGTC-----ATCAAGATG----- 1089  
QY 271 uThrSerValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluG1 291  
-----

Db 1090 -----GTCAGGCCATGCGCGCGGTACTGCCCAA 1121  
Qy 291 yValPheLysThrTyrSerPro-----LysCysAspLeuValLeuAspAla----- 306  
Db 1122 AACCTCCACGCGCAGCAGCCACCAGCAAGTTCGACTGGAGTGGGTGCGT 1181  
Qy 307 ----AlaHisGluLysSerThrGluSerSerGlyProLysArgGluGluMetG1 325  
Db 1182 ACTGTCGAGCGCCCTGCGCGGAGACGGGACACCCCGCGCGCGGAATCTCTC 1241  
Qy 325 uSerIleLeuPheLysCysSerAspPheValValGlnPheLysAspMetAspSerSe 345  
Db 1242 CTTGCGGTACGCGGACGACGACACGCTGCTCTGGAAACGACGCCCTTGAAGCGG 1301  
Qy 345 rTyrAlaLysArgAspAlaPheThrAspSerAlaLleSerAlaLysValAsnGlyLui 365  
Db 1302 TGCACCCGAAACACAGCGGACGCGCGCGCTCTCGGCTCGTGGCCACGCGC 1356  
Qy 365 sLysGluLysAspLeuProTrp----- 373  
Db 1357 -----GGGTAGTGGCGGTGTCTCGCCAAAGCTCTCGCGCGCTGCGCGC 1406  
Qy 374 -AspAlaGlyGluLeuThrAlaAsnGluGluLeuAlaLeuGluAsnAspValSerAs 393  
Db 1407 TCAGGCGAGCGTCTGTGAGCATCTGAGTCCGGGAGCGGACGCCAACCCGCTCATGT 1466  
Qy 393 nGlyTrpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerTh 413  
Db 1467 GGGCTGG----- 1473  
Qy 413 rTyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPh 433  
Db 1474 -----TCGCTGGCCACCCCGCGCGGTGGAGCAC----- 1506  
Qy 433 eLeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTy 453  
Db 1507 -----CGCGCGTCTCTCGGACGGATGCGAGGAGGATGCGGAC 1550  
Qy 453 rLysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluGluLysTyrThrAlaVa 473  
Db 1551 GCGCGGCTCTGCG-----GAGGGCGCGCTGACCGCTCTCTGTGTCACC----- 1596  
Qy 473 lGlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLys-TyrI 493  
Db 1597 -----GGACAGCCGG-AAACAGACGGCAAAACCGTGTCA 1630  
Qy 493 leProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGln-AsnSer 512  
Db 1631 TCTTCCCGCCCAAGCGCCCAATGGGTG-----GGCATGGGAGCCCAACTCTCTCA 1681  
Qy 513 ---ProArgMetGlyGlnPro----- 518  
Db 1682 ACACCTACCGCTTCGCGCGCGCTGCTGAGTGGCGGATGCTCTAGCGCGTATA 1741  
Qy 519 -----GlySerGlySerMetProSerArg----- 526  
Db 1742 CCGACTGGTGGCTCATCGAGTATCATCCGCGCAGCGCGCGCTCCCTCGCTTGAGCGTG 1801  
Qy 527 SerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnArgValValAsn 546  
Db 1802 TCGACGTCGTAC----- 1813  
Qy 547 GlyGlyValProTrpProSerProCysProSerProSerSerArgProProSerArgTyr 566  
Db 1814 -----AGCCCGCCACTTCGCGCTGCTGCTCC----- 1843  
Qy 567 GlnSerGlyProAsnSerLeuProArgAlaAlaThrProThrArgProProSerArg 586  
Db 1844 -----TCGCGGCACTCTGGCAATCCGTGGCATCTCCACCCGCGAG-----CGGTATCGGCC 1894  
Qy 587 ProProSerArgProSerArgProProSerHisProSerAlaHisGlySerProAlaPro 606  
Db 1895 ACTCCCAAGCGAAATCGCGCGCGCTGCGTGGCGGACACCTCACCCCTCACCAACGCGC 1954

Qy 607 ValSerThrMetProLys-ArgMetSerSerGluGlyProProArgMetSerProLys-- 625  
Db 1955 CCAAAATCGTCACCTCCGCGAGCCAGACATCGCCACCATCGCCGACACGGCGCA 2014  
Qy 626 ----AlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSe 644  
Db 2015 TGATGTCCTCGCCACCCCGCGACACCATCG-ACCTCACCACTGGCAGCGCAACTC 2073  
Qy 644 rGlyLeu---GluPheValSerHisAsnProProSerGluAlaAlaThrPro---ProVa 662  
Db 2074 TGGATCGCGCACACAAACAGCCCAACCGCTCATCGCAGGCGACACCGCGCTG 2133  
Qy 662 lAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValProArgLe 682  
Db 2134 CACCAACTCCACACCATACACCCAGGAGGACAGAGCCGCGCATCATCCCCGTGAC 2193  
Qy 682 uSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsn----- 699  
Db 2194 TAGGCTCCACACCGGACACGTCGACACCATCAAAACACCATCAACAGACGTACTCGAC 2253  
Qy 700 ----ThrProSerGlyProValLeuAla-----SerProGlnAlaGlyI1 713  
Db 2254 GGCATCACCTCGAACCCGCGCACCATCTCCCTGGCTCTCCACCGTCAACGAGCAGTGGATC 2313  
Qy 713 eIleProThrGluAla-----ValAlaMetProIleProAlaAlaSe 727  
Db 2314 GAACCAACACCGTCGGGACAGCTACTGGTACCGCAACCTCCGCAACCGTGTCAATTC 2373  
Qy 727 rProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLysAs 747  
Db 2374 GAGCACCATCTCCACACCTCCCGGACAGGCTACCGCACCTACATCGAAATCA----- 2428  
Qy 747 pSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnIleLysPr 767  
Db 2429 -----GCCCC 2433  
Qy 767 oAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerProValValSe 787  
Db 2434 CACCCCGTCTCACCACCGCATCCCAAGAAACCTCGAAGCCACACACACCCCAACA-- 2491  
Qy 787 rGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnPr 807  
Db 2492 -----CCACC 2496  
Qy 807 oSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyLysSe 827  
Db 2497 ATCGTACCGGCA----- 2509  
Qy 827 rArgAspLeuIleLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSe 847  
Db 2510 -----CCCTCCGCG----- 2518  
Qy 847 rSerSerAsnCysThrSerGlySerSerLysProAsnSerProSerSerIleSerProSerI1 867  
Db 2519 -GGAGCAGCAGACACCCCGCTCTCTACCAACCTCGCCACCTCCACCAACA----- 2572  
Qy 867 eLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSe 887  
Db 2573 -----ACGGAACAC-----CAGTCAAC 2589  
Qy 887 rSerProAlaCysLysGlnGluLysAspAspLysGluGluLysAspAlaAlaGluG1 907  
Db 2590 TGGCCCAACCTCTTCACAGGCA----- 2611  
Qy 907 nValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGl 927  
Db 2612 -----CCCAACCCA-----CCGCATCCCCCTCCCC 2637  
Qy 927 nProLysProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMe 947  
Db 2638 ACTACCCCTTCCAAACACCACTACTGTGCTCCCC---GCAACACACGCGGAGCGAT 2694

```
QY 947 t-----ValGlyHisGlnGlnProThrPro-----Va 956
Db 2695 GTGAGTGGCGTGGCGCTCCAGGGCCAGCGCCACCCTGGCGGGCGCGTGTGTGAGCGTG 2754
QY 956 lTyrThrGlnProValCysPheAlaPro---AsnMetMetTyrProValProValSerPr 975
Db 2755 CCGGACACGGGGGTGTGTCTACCGGGCAGTGTGTGGTGGCCA-----2800
QY 975 oGlyValGlnProLeuTyrProIleProMetThrProMetProValAsn-----G1 992
Db 2801 -----CCACCCGTGGCTGGCGGACCAACCGCGTCTCCGGAACGGTGTGTGTGGCG 2850
QY 992 nAlaLysThrTyrArgAlaValProAsnMetProGlnGlnArg-----1006
Db 2851 GGCACCGCATGGCGGAATCGCCATCGCGCGCGAGACGAGACCGACCCACCCCTG 2910
QY 1007 -----GlnAspG1 1009
Db 2911 GAAGAGCTGGTCATCGGCCAGCGATGACACTGCCGGAAGACGGTGCACTACATGTCAG 2970
QY 1009 nHisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaTh 1029
Db 2971 GTACTGGTGGCGGGTGGAGGACGGCGCGGAGGGTGGGATCTACTCTCGCCCCGAC 3030
QY 1029 rProPro-----AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPh 1045
Db 3031 GCGGCCAGGAACAGGAATGGCTGGAGCAGCGCTCGG---GCACACTCGCCACGACGGC 3087
QY 1045 eProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTy 1065
Db 3088 GACGGTTCGGCGCGGGCGGCA-----TGGAGAACGGCATGCCCGAGTGGCGG 3135
QY 1065 sSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProG1 1085
Db 3136 CCGCCCGGTGTCGAGCCGATCGCTCTGGATGACTTCTACGACGACCTCGCCAGCGCGG 3195
QY 1085 yLeuValSerSerAlaThrGlnTyrGlyAlaHisGlnGlnThrHisAlaMetTyrAl 1105
Db 3196 TATG---AGTACGGCGCCGCTTCGCGCGGCTGGAAGCGGTCTGGAAGCGCATGGCGAG 3252
QY 1105 aCys-----ProLysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAl 1122
Db 3253 GTGTTCGGGAGCGCGCTGCGGAGGAGCAGACGGAGCTCGCGGCG---GGTTC 3306
QY 1122 alleSer-ThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProH 1142
Db 3307 GGTATCCATCCGCGCTGTGGACGCGGTGGCACGCGGCAACTTCTGTGT-----3359
QY 1142 isThrProHisProGlnProSerAlaThrPro-----ThrGlyGlnGlnGlnSerGlnH 1160
Db 3360 --GCCCGCGCGCGCGGCGCAACGCTCTCCCTTCGTGTGGAACGGCGTACGGTGTGTG 3417
QY 1160 isGlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnA 1180
Db 3418 GCGGGGGAGGACACGCGCGTCCG-----TGTGCGCGCC 3450
QY 1180 laLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProT 1200
Db 3451 CCGCCACCG-GCACGGACTCGTTCCAGCATCAGCC-----TGTTCGACAGCACCGG 3500
QY 1200 hrProProSer-----MetThrPro---AlaSerAsnThrGlnSerProGlnAsnSerP 1217
Db 3501 CTCCCCCTGCGCTCGGTGGACTCCCTGTCTCCGGCGCATCAGTCCCGAGCAGCTCGC 3560
QY 1217 heProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrT 1237
Db 3561 TCCCG-----3565
QY 1237 hrAsnProProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProS 1257
Db 3566 -----CCTCCGCGGTGCGGTCGCTCGTGTACGCTGTGAGTGGAGTGGACCGA 3620
QY 1257 erHisProThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProG 1277
```

```
Db 3621 GCACCCCA-----||||| ||||| ||||| |||||
QY 1277 lnAlaAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThr 1292
Db 3645 GGTTCCTGGGCCACCCCTCGCGGATG---CCACACCCAGCAGTGGACG 3688
RESULT 7
US-11-075-185-37/c
; Sequence 37, Application US/11075185
; Publication No. US2005026643A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMERUTICINS
; FILE REFERENCE: 010099, 03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 37
; LENGTH: 10524
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-37
Alignment Scores:
Pred. No.: 3,21e-07 Length: 10524
Score: 416.00 Matches: 324
Percent Similarity: 32.94% Conservative: 155
Best Local Similarity: 22.28% Mismatches: 602
Query Match: 5.98% Indels: 375
DB: 7 Gaps: 61
US-10-802-228-2 (1-1312) x US-11-075-185-37 (1-10524)
QY 5 AlaAlaAlaProArgSer-----ProAlaValAlaThrGluSerArgArgPhe 20
Db 5186 GCCCGCGCGCGCGCAGCGAGCAGCGCGCCCTCGCCCTCGAGAGTTCGAGGACC 5127
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db 5126 GCCCGGAAT-----CGTTCGTCTTGGGGCCGTCGGCCAGGAGCGCGCC 5082
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 5081 AGCGCGAGGCTTTCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5022
QY 61 -----GlyProProProSerArgGlnSerSer 69
Db 5021 CGCACAGCTCGCAGCATCGCGAAGCGCGCGCGCGCGCGCGCGCGCTCGTTCGAGC 4962
QY 70 ProProSerAlaSerAspCysPheGlySerAsnGlyAsnGly-GlyGlyAlaPheArgPr 89
Db 4961 CCGCGGAGATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCC 4902
QY 89 oGlySerArgArgLeuLeuGlyLeu-----GlyGlyPr 100
Db 4901 AGCAGCTGAGCCCTTCTTGGCGGACATCGTCACGACACCCAGCGCTTCATCGGCG 4842
QY 100 oProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProPro-----118
Db 4841 ATGTGCGCGC-----CGCGAGGTGCGCGCTCATGTCG 4809
QY 119 -AlaAlaProThrArgAlaSerProLeu-----GlyAlaArgAlaSerProProArgSe 136
Db 4808 AGCCAGGCTCGCGCGCGAGACCCCATGCGAGGCTCATCGCGGCGACCGCGCGCGCG 4749
```



```
QY 136 rGlyValSerLeuAlaArg-----ProAlaProGl 146
:|||||
Dy 4748 AGGTGCGAGCGCGGTCCAGGAACGTGTTGGCCGCGCGGTAGTTCCGCTCGCCCGG 4689
146 yCysProArgProAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysProGl 166
:|||||
Dy 4688 CTGC-----CCAGGTCGCGCGCG-----CGGACGAGACACAGACAGAGCGCGCG 4644
166 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 186
:|||||
Dy 4643 AGCGCGCGGTGCTGCTGAGCTCGTGGAGATGAAGAGCGCGCTGACCTTCGCGCGCAGC 4584
186 nGlnProProAlaAlaAlaAsn----- 194
:|||||
Dy 4583 AGCGCGCGCATCTGCTCGCGCGAGAGCTTGGTCACGAGCTGTCTGTCAGCACACCTCGC 4524
195 ---ValArgLysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerPr 213
:|||||
Dy 4523 GTGTGCACAGCGCGCTCAGGGGAC-----GCGCGCGCGGATCGCTCCACACAC 4473
213 oSerSerSerValSerSerSerAlaThrAlaProSerSerValAlaAlaAlaTh 233
:|||||
Dy 4472 GCC-GCCAGCGCGCTGCTGCGCGCTGCGCGTCCGACGCGGCGAGCTCGACGTCGCGCGCGC 4414
233 rSerGlyGlyArgProGlyLeuGlyArgGlyArgGlySerAsnLysGlyLeuProGl 253
:|||||
Dy 4413 CGCG-----GCGCGAGCTGCTCCACCAGCGCGCGCGCTGCGCGCGCTCATCCCGCG 4360
253 nSerThrLysSerPheAspGlyLeuTyrAlaAsnMetArgMetValHisLeuThrSe 273
:|||||
Dy 4359 CGCGAGCTCAGC-----ACGAGTGGCGCGACCGC 4330
273 rValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPh 293
:|||||
Dy 4329 GTGCGCGCTCACCAGTGC-----TT 4309
293 eLysThrTyrSerProLysCysAspLeu-----ValLeuAspAlaAlaHisGluLysSe 311
:|||||
Dy 4308 CGCGACATCCCGCGAGCTCCCGGTGCGCGCGGTGATCAGCGCGCTCCCGTCCGCGGTC 4249
311 rThrGluSerSerSerGly---ProLysArgGluGluLeuMetGluSerIleLeuPheLys 330
:|||||
Dy 4248 CAGCGCGATCCGCGCGCGCGCGCGCGCGCGATCAGAGCACACCGAGC----- 4200
330 sCysSerAspPheValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAs 350
:|||||
Dy 4199 ---CGGGGCGCAGCGCGATCCCTGCGGAGCGCGCAGCTCGGGTCCGCGACGCTCTC 4144
350 pAlaPheThrAspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLe 370
:|||||
Dy 4143 GAGGGCGCTGTCGAGCAGCTCCGCGCAGAGCGGTTCAGGTCCGAGTTCGATGAGGCGCAC 4084
370 uGluProTrpAspAlaGlyLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAs 390
:|||||
Dy 4083 CCGCGCGTCCGGGTGCTACTCGCACTCGCGT-----AAGATCC 4042
390 pValSerAsnGlyTyrAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyVa 410
:|||||
Dy 4041 CCAGAGCGCGGTGCGCC-----AGGTCTTCACCCCGCGCGCGCGC 3997
410 lValSerThrTyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSe 430
:|||||
Dy 3996 GGTGCGACACCGCTCGCGGTCCAGCACGAGCTCGGTGCTCGTTCAGTTCGCGCGTCCGC 3937
430 rGluGluPheLeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluLleGluSerSe 450
:|||||
Dy 3936 CAGCCATCGCTGC---ACGAGCTCGAGGCGCTGCGCGCTGCTGTCAGACAGCGCGCAGG 3880
450 rAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspArgSerGluGluGluLysTyr 470
:|||||
Dy 3879 CAGCTCGCGCTCGCGCGGTGCTCGCGGTGTCACACAGAGCGCTGGGCGCGCGCGCGC 3820
470 rThrAlaValGlnArgAsnSerSerGluArgGlyGlyHisSerIleAsnThrArgGluAs 490
:|||||
```

Db 2904 GAAGACGACTGTCCCGACAGCCAGCGGTGTCTCCCTGTGCACAGCCGAC----- 2855  
Qy 798 ysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuL 818  
Db 2854 -----CGCCAGGATGTAGCTGTCTCGTCCGCGCCGACCCGGT 2818  
Qy 818 euAsnLysAsnArgGluGlyGluLysSerA-ArgAspLeuLeuLysAspLysleGluProS 838  
Db 2817 TCGCGCGCCAGACAGCGGCGCGCCGCTG-----CGC 2785  
Qy 838 erAlaLysAspSerPheIleGluAenSerSerSerAsnCysThrSerGlySerSerLysP 858  
Db 2784 GAGCCGA-----GGTCGCTCGGTGATCCGCTCCTCGACGC 2746  
Qy 858 roAsnSer-----ProSerIleSerProSerIleLeuSerA 870  
Db 2745 CTCGAGCCAGTAGCTGCCGTGGAACGCGTACGTGCGCAAGTCCACAGGTTGCCGCC 2686  
Qy 870 snThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProA 890  
Db 2685 GTACGGCGCGAGCACCCCTGGGCCCACTCCACCGAATGCCCTGGACGTGCAGCGGCCAG 2626  
Qy 890 laCysLysGlnGluLysAspAspLysGluLys-----LysA 903  
Db 2625 CGCTGAALAAAGACCCCGCCCGTCTGCGCTGCAAGTGCACCGCCACACAGCC 2566  
Qy 903 spAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnProA 923  
Db 2565 GCCGCGCTCGGCGCTCGCGCGCTCAGCGGATGCCAGCACCGGATCGCGCTGACCTC 2506  
Qy 923 rgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGlnAlaGlnP 943  
Db 2505 GACGAACACCCCGTCCCGCTCG-----CCAGACGCGGAGCGCGCGCTCGAG 2455  
Qy 943 roSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnProValCysP 963  
Db 2454 CGCACCGGCTCGCGAGTTCGCGCACAGTATGCCG----- 2417  
Qy 963 heAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeuTyrProI 983  
Db 2416 -----CGCGAGCTCGCGCGCGCGGAGCACTTCTCCGT 2383  
Qy 983 leProMetThr-----ProMetProValAsnGlnA 993  
Db 2382 CACCGTCGAGTACAGCGGGATGGCGACGACGCGGCTCGACCGGCGGCGAGCGCGCTC 2323  
Qy 993 laLysThrTyrArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHisGlnS 1013  
Db 2322 GAGCTCCGCGAGGATCGCTCCACGTGCGCTGTGCAGCGGTAAATCGACGTTCACCGC 2263  
Qy 1013 erAlaMetMetHisProAlaSerAlaAlaGlyPro-----ProIle-----A 1027  
Db 2262 CCGCGGAAGACCTCTCTGTCTCCAGCTCGCGCCAGCAAGCGATCGACCGTGTGCGCGCTC 2203  
Qy 1027 laAlaThrProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPhePro 1046  
Db 2202 CCGGACACCAACCG-----TCGAGCCGCGGTGTTCACCGCGCGGCGGAGCGTGC 2149  
Qy 1047 Asn-----GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHis 1063  
Db 2148 TCCGNAGCGAGAGCCCGTCTCTC-----CACCTCGCGGACCGGGCGCTCGATCAGGCCAT 2092  
Qy 1064 -ValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGl 1083  
Db 2091 CGCGCGCGCGCGGATCGCGCGGACGAGCTGGCTGCGGAGCGGCGACGCGCGCGC 2032  
Qy 1083 nProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMe 1103  
Db 2031 GTCCTCGAGCGTCAAGCGCCCGCGGACG----- 2004  
Qy 1103 tTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaI 1123  
Db 2003 -----TAGCGCGCGGACCTCGCCCTGGCTGTGCGCGCACGAC 1966

Qy 1123 eserThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAla-ThrLeuHisProHisT 1143  
Db 1965 GCGCAAGGCTCGATG-----CCCCACGCCCGCAGACCGCGGAG 1924  
Qy 1143 hrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlyS 1163  
Db 1923 CCCACGCGCATCGCAACAGCG---CGGCTGCACCAACGTCACGCGGTCCGCGGGCGG 1867  
Qy 1163 erHisProAlaProSer-ProValGlnHisHisGlnHisGlnAlaAlaGlnAlaLeuHis 1182  
Db 1866 CACGCGCTCTCTCTGCTCGCGCGGAGGACCGAGGACCGACCGCCGCT----- 1816  
Qy 1183 LeuAlaSerProGlnGlnSerAla-----IleTyrHisAlaGlyLeu 1197  
Db 1815 --GAACGCGCGACGCGCGGTCTCGAGCGCGCGACCGCGCGGCAAAACACTCGGACTC 1759  
Qy 1198 AlaProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhe 1217  
Db 1758 GG-CCAGCAGCGCGCGCCCATCGCGGCACTGGCTCCCTGGCCCGGAG----- 1709  
Qy 1218 ProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThr 1237  
Db 1708 CCAGAACACGA----- 1697  
Qy 1238 AsnProProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSer 1257  
Db 1696 -----CCTTCGCGCGCGCGCGCTCGCGCGGTTCACCGGGGTGCGGCTCCCTT 1643  
Qy 1258 HisProThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGln 1277  
Db 1642 GCGAGAGCGCGCTCAGGCGCGCGCGCGCTCGCGCGGCTCGCGCGGTTCACCGCGCGC 1583  
Qy 1278 -----AlaAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThr 1293  
Db 1582 GCGCTCCAGTGTGTCGCGCGCGCGCGCGGTGTGCACCGCTCGCGCGCGCT 1523  
Qy 1294 AlaHisPheProTyrMetThrHisProSer 1303  
Db 1522 CCGGTGCGCGCGAGCGCACCGCGCGCGC 1493

## RESULT 8

US-11-121-086-4  
; Sequence 4, Application US/1121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POULSEN, TIM S.  
; APPLICANT: NIELSEN, KIRSTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121.086  
; PRIOR FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 4  
; LENGTH: 164810  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-4

## Alignment Scores:

Pred. No.:	7,54e-06	Length:	164810
Score:	407.00	Matches:	384
Percent Similarity:	31.84%	Conservative:	156
Best Local Similarity:	22.84%	Mismatches:	581
Query Match:	5.85%	Indels:	583
DB:		Gaps:	78

US-10-802-228-2 (1-1312) x US-11-121-086-4 (1-164810)

Qy 2 ArgSerAlaAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgPhe--- 20



QY 558 roSerSerArgPro-----ProSerArgTyrGlnSerGlyP 570  
DB 87313 CCGGCCCGGCCAGGAGGACGACTCTGCACTGACTCCGCGGAGGACGCGCGTCAA 87372  
QY 570 roAsnSerLeuProProArgAlaAlaThr----- 579  
DB 87373 CTTCTTAGCGCGCTCGCTCAGCGAGGAGAGCGTGAGGGACCCCGAGGGACCGC 87432  
QY 580 -----ProThrArgProProProSerArgProProSerArgP 594  
DB 87433 CGGGGACACGCGGCGGCAAGACCCACCGCGCGCCCACT---CTGCAGACAGCG 87489  
QY 594 roProSerHisProSerAlaHisGlySerProAlaProVal---SerThrMetProLysA 613  
DB 87490 CATGGGGCACCGGCACAAAGCGGGAGGCGCGCGCGGAGCGAGTCTCGGGGGCGC 87549  
QY 613 rgMetSerSerGlu--- 620  
DB 87550 GGGCAAGAACAGACGAGGCGTGGAGCTGCCCTGGCGCGCGCCCGCGAGCGCGCCGCGAGA 87609  
QY 621 -----ArgMetSerProLysAlaG 627  
DB 87610 CAAGGACGGCTCAAGCCCGCGGACCGCGGGGACCGGGGCGCTCCAGTCTGCTGTGCT 87669  
QY 627 lnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuG 647  
DB 87670 CACGACGCGCCACTGAGGAGGCGGTGTACTGCTTCTAC--GGCAACGACTCG----- 87718  
QY 647 luPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArgThrSer- 666  
DB 87719 -----GACGAGAGCCCGCGGCGCGGCCACCGCCACCAACCCACCGCGGCGCATCGG 87770  
QY 667 ProSerGlyGlyThrTrpSerSerValValSerGlyValProArgLeuSerProLysThr 686  
DB 87771 CCATCC-----CTCGCGCTTTACGCGGAGCGTCCGCGAGGCGCGGAGGAGG 87818  
QY 687 HisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyPro----- 704  
DB 87819 CCCCTGCGCCGTF---CCAAGGTGACCAAGTCCCGCGCGCGCGCGCGCGGACCCAGCCCA 87875  
QY 705 -----ValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetPro 722  
DB 87876 GCCTCATGTGACAGACCCGCGCTGCT-----ACTCCGTGAGCTCTCCGCGCA 87926  
QY 723 lleProAlaAlaSerProThrProAlaSer-----ProAsnSerAsn 736  
DB 87927 GCTCCCTCAGCGAGCCCGAGCCCTCGAGCCCGCGCGCGCTCCATCCAGAGCGCGGAGC 87986  
QY 737 ArgAlaValThr-----ProSerSerGlu 744  
DB 87987 CC-GGGGTACCAAGGACCCGGGCGCCAGGAGCGGACGACGACGCTCGCCCGACCGCGG 88045  
QY 745 AlalysAspSerArgLeuGlnAsp-----GlnArgGlnAsnSerPro 758  
DB 88046 GCCCGGAGGAGCTTCTGCAGCGGTGCATACAGTCCGCCCTGCGCGCGCGCGCGCC 88105  
QY 759 AlaGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGlu 778  
DB 88106 GTGTCTGCGCTCGCGCGCGCAAGCCCGGAGCCACCCGCGTGGATGAGCGCGCGCGCAG 88165  
QY 779 -----AsnLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAsp 795  
DB 88166 GGGTCCCGGGAACGCGGCGGAGGAGCAGCGGCTCGGACCGGCGCTCCGACCTGTGATAGC 88225  
QY 796 -----LeuLysLysPheLysAsnAspPhe----- 803  
DB 88226 GTGGAGTGGCGGCATCCAGAGGGCGCCCAATTGCAATTGTCACGTGCTGCACCCAGGCA 88285  
QY 804 -----ArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsn--- 819  
DB 88286 GCAGTGCACCGCGGAGGCGCTCGTCCGAGTCCGACTCCATCTCTCTCTCTCTCTCTCT 88345

QY 820 -----LysAsnArgGluGluGlyLysSerArgAsp 829  
DB 88346 CTGTCAGTGGGATCACCTACAGCCCCCAAGCACAGAAAGGACGACGAGCGGAGGA 88405  
QY 830 LeuIleLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSerSerSer 849  
DB 88406 GAAATGGGCAGTCCCGCGCGCCAGAGAAAGG-----GGCGCACCTCA 88450  
QY 850 AsnCysThrSerGlySerSerLys----- 857  
DB 88451 GTCAGACACGCGGAGCGCCCGTTCCTCGACGCCCGAGAGCCAGCTGGCACACAG 88510  
QY 858 -----ProAsnSerProSerIle-----Ser 864  
DB 88511 AAGACACGCGCGGGTGCAGCTGTCTCCGGGACGAACAGTGTATCTACCTCCCCAGC 88570  
QY 865 ProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGlnGlyVal 884  
DB 88571 CCGCACCCCTGCCCCACGCCCAAGAGGACCCCGCGCGCCCGCCACACCGCGGAAGGTG 88630  
QY 885 -----GlnThrSerSerProAlaCysLysGlnGluLysAspAspLys 898  
DB 88631 CGGCCCTTCTGCTGGCACAGCCCGCGCTCAGCC----- 88666  
QY 899 GluGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLys 918  
DB 88667 -----AAAGTCCGAGCGCGCGGCGAGCGGTCTCGGAGCGCTACACCGCGCTGCCAAG 88720  
QY 919 GluPheAsnProArgSerPheSerGlnProLysProSerThrThrPro----- 934  
DB 88721 ACCTCGAGGTGGCGAGCTGAGCCAGCCCGCCAGAGCCACACCGCGCGCGCTC 88780  
QY 935 -----ThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGln 952  
DB 88781 GCCAAGACCCCTCTCCAGCTCTCCAGACTCGCGCGC-----TCCAG 88828  
QY 953 ProThrProValThrThrGlnProValCysPheAlaProAsnMetMetTyrProValPro 972  
DB 88829 CCCCTGCCAGAAAGCGCGCGCGTCAACCGAGTGTCTGGCGCGCTG----- 88876  
QY 973 ValSerProGlyValGlnProLeuTyrProIleProMetThrProMetProValAsnGln 992  
DB 88877 CCGCGCGCGGAGCTCCCGGTGCCAAAGCGCGCGCGCACCTCTCTCGCGAGCAG 88936  
QY 993 AlalysThrTyrArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHisGln 1012  
DB 88937 CACAAGACGACAGATCGCCCGTGGATCCG----- 88969  
QY 1013 SerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThrProProAla 1032  
DB 88970 ---TTCATGCAGAGCGCGCGCGTGGCGCGCACCGCTGGCTCGGCGAGTCCCGGAG 89026  
QY 1033 TyrSerThrGln----- 1036  
DB 89027 CCGGGCCCCAGGGCGCGGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCTG 89086  
QY 1037 -----TyrValAlaTyrSerProGlnGlnPheProAsn-----GlnProLeuVal 1051  
DB 89087 GGCCTGGTGGTGTGGCTC-AGCCCTCTCAGCGCGGACGAGTCTCCGACCGCTC- 89142  
QY 1052 GlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGly 1071  
DB 89143 ---GGGCTTCCGGCGACAGCTAACCTTCAT-----CAAGGA 89175  
QY 1072 AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal-SerSerSerAl 1091  
DB 89176 GTCGCGGGCTGGCGCGCGCGCTCGAGCTGTCTCGCGCGAGTCCGCGCGCTCTGC 89235  
QY 1091 aThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLysLeuProTy 1111  
DB 89236 CCCCAG---GGCGCTCGCGCGCGCG- CGGCGCGCGCGCGCTGCCCGCGCTCTCTCTCT 89291  
QY 1111 rAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnG 1131

Db 89292 G-----CTCCTCGCGCTGG 89306  
 QY 1131 nTyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaTh 1151  
 Db 89307 NAGAGCT---CCGAGCGGCACC---CCGCAGGGCCCGCCCGCCCGCGCAGCGC 89357  
 QY 1151 rProThrGlyGlnGlnSerGlnHisGlySerHisProAlaProSerProValG1 1171  
 Db 89358 CCCCCGCGCGCCGACCCAGCCCTGGCGA-----GCGCCCTGCCCGCGCGCA 89402  
 QY 1171 nHisHisGlnHisGlnAla-----AL 1178  
 Db 89403 CCACCTCCGAGAGCCCGTCCCGCTGCTGCGCGCGCCCGCCCGCGCGAGACTG 89462  
 QY 1178 aGlnAlaLeuHisLeuAlaSerProGlnGlnSerAlaLalleTyrHisAlaGlyLeuAl 1198  
 Db 89463 TCAGCGCTACCGTCTGCTGCCGACATCAGCGTGGCCCG-----CAGGCCCG-ACGCG 89515  
 QY 1198 aProThrPro-----ProSerMetThrPro-----AlaSerAsnThrGlnSerPr 1213  
 Db 89516 GCGTCTCCCGCGCCCTGCTCCTCAGCGGAGCGCGCGCGCGCGCAGCAGCGCGAGCCC 89575  
 QY 1213 oGlnHisSer-----1216  
 Db 89576 CGCGCGCTCCCGAGGTGGCGCGCGCGCGCGACCTGGCGCGCATCCGAGTAGAGGAC 89635  
 QY 1217 -PheProAlaAlaGlnGlnThrValPheThrThrHisProSerHisVal-----G1 1233  
 Db 89636 GTGCCCCACATCTCGCGAGCAGCTTCCGCCACGCGCCCTGCCACTCGGGGCTCCACG 89695  
 QY 1233 nProAlaTyrThrAsnPro---ProHisMetAlaHis-----ValProGlnAlaHisVa 1250  
 Db 89696 CCGAGGAGCGCCCGCGCGCGCCCGCGCGCGCAAGCAGCAGCGCGCTGGTGCAGACC 89755  
 QY 1250 lGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMetThrTh 1270  
 Db 89756 GAGGAGGTGCG---CCGCGCCCAAGACCACTCCAGCAGCTGCCCGAGCGTGGAGACAGG 89812  
 QY 1270 rGlnProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProleProVa 1290  
 Db 89813 GAGCCCCCGGGCGCCCGCGCGCGCGCGAGCTCTCCCTCTCGGAGCAGCGTGGACGCT 89872  
 QY 1290 lSerThrThrAlaHisPheProTyrMetThrHisProSer 1303  
 Db 89873 CCGAGCTCGCCCAAGGCTCCCA---TCTCGCACCTCTCG 89909

## RESULT 9

US-11-075-185-34/c  
 ; Sequence 34, Application US/11075185  
 ; Publication No. US20050266434A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: REEVES, CHRISTOPHER D  
 ; APPLICANT: JULIEN, BRYAN  
 ; APPLICANT: REID, RALPH  
 ; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS  
 ; FILE REFERENCE: 010099.03  
 ; CURRENT APPLICATION NUMBER: US/11/075,185  
 ; CURRENT FILING DATE: 2005-03-07  
 ; PRIOR APPLICATION NUMBER: US 60/551,103  
 ; PRIOR FILING DATE: 2004-03-08  
 ; PRIOR APPLICATION NUMBER: US 60/568,290  
 ; PRIOR FILING DATE: 2004-05-04  
 ; NUMBER OF SEQ ID NOS: 61  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 34  
 ; LENGTH: 11070  
 ; TYPE: DNA  
 ; ORGANISM: Sorangium cellulosum  
 US-11-075-185-34

## Alignment Scores:

Pred. No.: 1.54e-06 Length: 11070

Score: 399.00 Matches: 343  
 Percent Similarity: 33.55% Conservative: 167  
 Best Local Similarity: 22.57% Mismatches: 553  
 Query Match: 5.73% Indels: 462  
 DB: 7 Gaps: 69  
 US-10-802-228-2 (1-1312) x US-11-075-185-34 (1-11070)  
 QY 18 ArgArgPheAlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArg 37  
 Db 6253 CQTCAGGCGCGCGCGGAGAGCGGGTGGTGGCATACAGGCGCGCCCATGCGCGGCG 6194  
 QY 38 SerGlyArgGlyGlyGlyAlaAlaProGlyProTyrPro-----51  
 Db 6193 GCTGCTCCCTGGCGGTGAGACACGCGGACCTTGCCTGTCTCTCGACGGGG 6134  
 QY 52 SerAlaAlaPro-----ProPro 57  
 Db 6133 TCTGCGATCGCTCGTCGAGCGCTCTCCCTGGCAACGCTCCAGCTCATCATCC 6074  
 QY 58 ProGlyProGlyProProProSerArgGlnSerProProSerAlaSerAsp---Cys 76  
 Db 6073 ACCCATCCGGGGCGCCCGCGAGAGCGGCGAGCGCGCTGGCGAGATGCTGTC 6014  
 QY 77 PheGlySerAsnGlyAsnGlyGlyAlaPheArgProGlySerArgLeuLeuGly 96  
 Db 6013 CGTGTGTGCAAGGCTCG-----5996  
 QY 97 LeuGlyGlyProProArgProPheValValLeuLeuProLeuAlaSerProGlyAla 116  
 Db 5995 -----CTGCCAGTCCA-----GCAGCCGCTCGTCCGGGTGCG 5963  
 QY 117 -----ProProAlaAlaProThrArgAlaSerPro-----126  
 Db 5962 ACCGCAAGTGTCCACCAACCGCGCACCTCGCGCGCAGCGCGCTCGTCCGCCCGCG 5903  
 QY 127 -----LeuGlyAlaArgAlaSerProProArgSerGlyValSerLeuAlaArg 142  
 Db 5902 AGACCAGCAGCGCGAGGATCGGAAGCGCTCCACCGCTTCGCGTCCGCGGGCG---5846  
 QY 143 ProAlaProGlyCysProArg-----149  
 Db 5845 CCGGTCTCGCGCGCTCTCTCGAGTACGTGCGCGTGTGTGCGTATGCCAAGGACG 5786  
 QY 150 -----Pro-AlaCysGluProValTyrGly-----Probe 159  
 Db 5785 AGACCGCGCGCGCGCGCGTCTCGTCCGCGCCACGCGCGGCTCGTTCAGGAGCT 5726  
 QY 159 uThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 179  
 Db 5725 TGACGTGCCCTCGCACCATGTCAGCTGCGCGGAGGGCGGTGCGCGTGCAGCGTCTTCG 5666  
 QY 179 nGln-----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 194  
 Db 5665 GAGCTCTCGGTGCTGCATCGCCAGCACCATGTTATCAGCCCGCCACACCGCCCGCGCA 5606  
 QY 194 nValArgLysPro-----198  
 Db 5605 CTGCGTGTGCCCGAGGTGGACTTGAGCGAGCCGAGCCAGCGCGCTCGCGCGGTGT 5546  
 QY 199 -----GlyGlySerGly-----Leu-LeuAlaSerProAlaAlaAlap 211  
 Db 5545 GCGCGGTGCGTAGTGTGCGCAGCGCGCTCGATGCGCTCGATGGGGTTCACCGAGCGGTTC 5486  
 QY 211 roSerProSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerValVal 231  
 Db 5485 CCGTCCGCTCGCGCTCGAGCGGTCTGATGTCCCGCGGGTTCAGCCCGCGCC-----G 5435  
 QY 231 laAlaThrSer-GlyGlyGlyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGly 250  
 Db 5434 CGGCCAGGGCTGGCGGATGACGTCTCTG--CTGGGCGCGCGCGCTTGGGGCGGTGAGCCC 5377  
 QY 251 LeuProGlnSerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIle 270

Db 5376 CTG---GTCGCGCGCTCTGCTTACGCGCGAGCTGCGGATGACCGCAG---CACGCG 5323  
Qy 271 LeuThrSerValValGlySerLysCysGluValGlnValLysAsn----- 285  
Db 5322 GTCGCGCTCGCGCGCGCTCGCAGCGCGCTTACGCGCAGGATCCGCGACGCCCTCCGA 5263  
Qy 286 -----GlyGlyIleTyrgluGlyValPheLysThrTyrsrProLysCys 300  
Db 5262 CCAGCGCGCGCATCGCGCGCGCGCGAAGCGCTTGCAACGACGCTCCAGCGCGTGCC 5203  
Qy 301 AspLeuVal-----LeuAspAlaAlaHisGluLysSerThrGluSerSer----- 315  
Db 5202 GCGCTGCGCGCTGAACATCGACGAAGATCATCGCGCTGGCCATCACGGTGGCGCGCGCC 5143  
Qy 316 -----SerGlyProLysArgGluGluIleMetGluSerIleLeuPhe 329  
Db 5142 GAGGCCAGGTGCGACTCGCGCGCGCGGAGCACATGCA----- 5104  
Qy 330 LysCysSerAspPheValValGlnPheLysAspMetAspSerSerTyrAlaLysArg 349  
Db 5103 -----CGCGAGGTGGACCGCGCACGAGCGACGA 5074  
Qy 350 AspAlaPheThrAsp-----SerAlaIleSerAlaLysValAsnGlyGluHisLys 366  
Db 5073 GCACGCGGTGTCGACGGTGACCGCGCGCGCGAGCCCCAGCGTGTAGCGCATCGCGCC 5014  
Qy 367 GluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAlaAsnGluGluLeuAla 386  
Db 5013 CGAGCGGACGCTGAGGAA---GTTCCGCTGAGGACGTGGCGCTCGAAGGACTCGAGCTG 4957  
Qy 387 LeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyrAsnGluGlu 406  
Db 4956 CTG-----CAAGAG 4948  
Qy 407 AsnTyrGlyValValSerThrTyrAspSerSerLeuSerSerTyrThrValProLeuGlu 426  
Db 4947 CGCGCGCGCGTAGTCCCGTGACGAGCGCGGAGAGCGCGCGTGGCGCTCGCCTCGAG 4888  
Qy 427 ArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla---AsnGlnLeuAlaGlu 445  
Db 4887 CGC-----GTGGGGCGCGAGGCCCGCGCGCTCGAGCGCTCCAGCGCTCCAGCGC 4846  
Qy 446 GluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspAspArgSer 465  
Db 4845 GCACCTCGAGCAGCAGCGCTGCTCGCGGTCCATCGCTCCGCTTCGCGCGGCTCATGCC 4786  
Qy 466 GluGluGluLysTyrThrAlaVal-----GlnArgAsnSerSer 478  
Db 4785 GAAGAAAGTTCGATCGAAGCGGATCGCGCTGCTAGGAAAGCGCGCGCTGCTGATCGA 4726  
Qy 479 GluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProPro----- 495  
Db 4725 CTTGCGCGGGGCTCGGGTGGGTGTACAGCCCGCGCACATCCAGCGCGCGCCCTC 4666  
Qy 496 GlyGlnArgAsnArgLuuValIleSerTrpGlySerGlyArgGlnAsnSerProArgMet 515  
Db 4665 GGGGAACGGCC-CGA-----TCGCGTCCGCGCGCTCCGAGACGA 4628  
Qy 516 Gly-----GlnProGlySerGlySerMetProSerArg 526  
Db 4627 GGTCCCCACAGCGCGCGGGGTATCGACGCGCGCGGAGCGCGGACGCCATCGACACGA 4568  
Qy 527 SerThrSer----- 529  
Db 4567 TCGCATCGCTGCTGCTGCGCTCCCGCGCGCGCGGGGCTGTGCGCGCTCGCGCCCGC 4508  
Qy 530 -----HisThrSerAspPheAsnProAsnSer----- 538  
Db 4507 CGAGCTCCGAGAACGACGCTGAAGCAGCAGATCCGCGATCGCGCGCGGTTCGATGTT 4448  
Qy 539 -----GlySer-----AspGlnArgValValAsnGlyGlyValProTrp 551  
Db 5399 -----GlySer-----AspGlnArgValValAsnGlyGlyValProTrp 551

Db 4447 CGAACGCGAGGGTCTGTCGGCAGCACCGTCTCGGCGCGCGCGCGCGT----- 4397  
Qy 552 ProSerProCysProSerProSerArgProProSerArgTyrGlnSerGlyProAsn 571  
Db 4396 -----TCTTACGCTCGAGCGCGCTGAGCGAATCGAGCCCGACCTCTTCA 4352  
Qy 572 SerLeuProProArgAlaAlaThrPro-----ThrArgProProSer 585  
Db 4351 GCGGCTTTCGCGCGGACCTGCGCGCGACCTGCGCGCGCAGCGCACGACCGCACCTCGC 4292  
Qy 586 ArgProProSerArgProProSerArgProProSerHisProSerAlaHisGlySerProAla 605  
Db 4291 CGCGCACGAGCTCGAGCAGCGCGCGCGCTCGCGCTCGGGCA-----ACGCGGAGA 4238  
Qy 606 ProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerProLys 625  
Db 4237 GCCGTGCGCGCAGCGCGCGCGG-----CCGCTGCTCGGCTTCGG 4196  
Qy 626 AlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSerGly 645  
Db 4195 GGGTCTGCGCTCGCG-----TCGGCGCGCGCGCGG----- 4166  
Qy 646 LeuGluPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArgThr 665  
Db 4165 -----TCGCGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCTCGCGCTCGC 4118  
Qy 666 SerProSerGlyGlyThrTrpSerSerValValSerGlyValProArgLeuSerProLys 684  
Db 4117 TCGCGCGACGCTGAGCG----- 4100  
Qy 686 ThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyProVal 705  
Db 4099 -----CCGACAGGTTCGAGCGCGCGCGCGCACCA--- 4073  
Qy 706 LeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetProIle---Pro 724  
Db 4072 ---GCGTCTGCTCGCGGTGCGCGCGCGCGCTCGAGGAGCGCGCGCTTCGCGCGACCG 4016  
Qy 725 AlaAlaSerProThrProAla-----SerProAlaSerAsnArgAlaValThrPro 741  
Db 4015 ACAGCGGACAAAGGCCCTGCGCTCGAGGCGTGCAGGTGCGCGCGCGCGAGTTCGCGC 3956  
Qy 742 SerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsn 761  
Db 3955 TCATGCGCGAGCGCGCTGCTCCAGAAAGCCCGAGGCTGCTCGCGCGCACCGCGC 3896  
Qy 762 LysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGly 781  
Db 3895 GCGCGCGAGCTGCGCGC---CGAACGCGTCCAGGAACGCAATTGGCAGCCCGCTAGTTTCG 3839  
Qy 782 IleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsn 801  
Db 3838 CTTGCGCGCGCTTCCCGACGCTACGAGATCGAG-GAGAACAGCACGACGAGCGCGCAGC 3780  
Qy 802 AspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsn 821  
Db 3779 GGCTGGTCTCTCGCGCGCGCTACACGTTGGCAGCGCGCTCGACCTC----- 3732  
Qy 822 ArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAsp 841  
Db 3731 ---GCGCGCAGGGCGCG-----GTGAGCTGCTCGCGCTCGC 3696  
Qy 842 SerPheIleGluAsnSerSerSerAsn-----CysThrSerGly----- 854  
Db 3695 GCCGTGAGCACCCGTCTCGCGCAGCACCCCGCGGTGTGCACACGCGCGCTCAGCGCGCGC 3636  
Qy 855 -----SerLysProAsnSerProSerIleSerProSerIle 867  
Db 3635 TTCGCGCGCGGACCTGGACTGCTGCCGAAAGCGCGCGCGCTCGTCACTCGCACGCC 3576  
Qy 868 LeuSerAsnThr-GluHisLysArg-GlyProGluValThrSerGln-----GlyValG 885  
Db 3575 ACGACCTCGACCGTCTCGCGCGCGCTCGCGCGCGAGCTGTTTCGACGAGCGCTCGCGCTCC 3516





```
Db 137098 CGGGCGCGCCCGCTCCGGCGCGGACCTCGGGCGCGGAGAGTCAAGGCGGCGCGG 137039
QY 87 Phe-----ArgProGlySerArg-----ArgLeuLeuGlyLeuGly 99
Db 137038 GCGGGCGCGCGGGTCCGCCCGCGGCGAGACCCCGCGCGGACTGTCGGAGCGCGG 136979
QY 100 ProProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAla 119
Db 136978 CGCGCGCGCGCCAC-----CGCCCGGGAGGGCGCGGCG 136943
QY 120 Ala-----ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArg--- 135
Db 136942 TCGGCGCTTTCGCGCGGAGCTCGGA---GGCGCGGGCGCCCTCCCGCCGTG 136886
QY 136-----SerGlyValSerLeuAlaArgProAlaProGlyCysPro 148
Db 136885 CCCCGCGCGGGGACTGGCGGGGTCTCCAGCTCGCGCGCCCTGCC----- 136832
QY 149 ArgProAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysProGlnGln 168
Db 136831 ---CGGCGC---CACCCGCTCTCAGGC----- 136811
QY 169 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPro 188
Db 136810 -----TCGTCCCGGGGCTGCCCGCAGCGACCCCGATTCGCGCG 136769
QY 189 ProProAlaAlaAlaAsnValArgLysProGly-----GlySerGlyLeuLeu--- 204
Db 136768 CCCCCCGCGGATCCGGGTGCGCGCGCTCGGGTCTCCGAGGGCGGGCGCACCCCTG 136709
QY 205 -----AlaSerProAlaAlaAlaProSerProSerSerSer 216
Db 136708 CTTCCCGCGCCCGCCACCGCGCTCCCGCGCCCTGGGTCCCGGGCTCTCACCTGC 136649
QY 217 SerValSerSerSerAlaThrAlaProSerValValAlaAlaThrSerGlyGly 236
Db 136648 CCGGTTCCAGCGCTATCGCAGCTTCCCAACCCCGCAGCGCGTCCCAATCGGGCTC 136589
QY 237 GlyArg-----ProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
Db 136588 GGCCAACTCCAGTCTCCCTCGGGGCGGGGCGAGATTTTCAGAGGGCGCTGGCCAAAG 136529
QY 254 Ser-----ThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetVal 268
Db 136528 TCCAAAGTGTCTGGGCTCGACCTCCCTGGGGCCCTCTCTGGGAGTCGCTCTCCGGACC 136469
QY 269 HisIleLeuThr-----Ser 273
Db 136468 CACGCTGTTCCTGGGGATGGAGTAAGTCCCTACACCTCAGCCCTTAAGGTGCCGGGAGC 136409
QY 274 ValValGly-SerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPh 293
Db 136408 ATAGTAGTGGACCATGTGGAGTGTGACAGCTGGGATGGATCCAAAGAGCC----- 136354
QY 293 eLysThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrG1 313
Db 136353 -CAGAGTGAAGCCCG---GCCGACAGCAAAATCAGTCTCCAGCGACGAATCTCGCAGGCA 136298
QY 313 uSerSerGlyProLysArgGluGluIleMetGlu-----SerIleLeuPhe-- 329
Db 136297 GGGGAGGCGCGTCCCGGGGCGACAGCTTGTGCAAAAGACCCGGGTCTGATCTGTTTCA 136238
QY 329 ----- 329
Db 136237 GGAACCGCAGTGATTTCTCTCTCTGTTAGGTGACACAGGTGGGAACTGGGGCCCTGG 136178
QY 329 ----- 329
Db 136177 CCTCTGCTTTGTCTTTTTCATCTATAGACAGAGTGTCTGGCCATGTTGGCCACGCTG 136118
QY 330 -----LysCysSerAspPh 334
Db 136117 GTCTTGAACCTCTGGGCTCAGCGGATCTTCCCACTGGGCTTCCCAAAAGTCTAGATTA 136058
```

```
QY 334 e-----ValValValGlnPh 339
Db 136057 CAGCGTGAAGCACTGCACCTGGCGTCCCTCTCTGTTTTTGCCCTTATTCATGCCACGCC 135998
QY 339 eLysAspMetAspSerSerTyrAlaLys----- 348
Db 135997 TCCCTCACCAGACTCTGAGCTTGGAGGTAGGCTGTCAAAACGCTGGCTGGACTTT 135948
QY 349 -----ArgAspAlaPheThrAsp-----SerAl 356
Db 135937 GCTGGCGATGCCAGATCGTGTGACGTCTCTGAATTAGTGTCAATCACCAGGCCAGT 135878
QY 356 aIleSerAlaLysValAsnGlyGluHisLysGlu----- 367
Db 135877 GCAGGAGAGGAACCTGCAGGCAAGCACCAACAGCCCTTGACCCCAAGGTCCCTGTGTG 135818
QY 368 -----LysAspLeuGluProTrpAspAlaGlyGluLeuThrAl 380
Db 135817 AGGCATTACCCAAACTCCTGCAGGGACCGTGAACCCAGAGATGGCAAGCTAACCTGTG 135758
QY 380 aAsnGluGluLeuAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMe 400
Db 135757 CCCCAGCAAGTG-----GGTGGGACCGGACCCACC 135725
QY 400 tPheArgTyrAsnGluAsnTyr-----GlyValValSerThrTyrAspSerSe 417
Db 135724 ACCGTCCACAGCTCTGAGACGTGGTGGCTCTGTCTGTCTCTCTGCGTCCCTCTCTC 135665
QY 417 rLeuSerSer-----TyrThrValProLeuGluArgAspAsnSe 430
Db 135664 CCCGATTCCCCAGCGGCGCACCGGCGCACACCACTGACCTTTGTTGAGGCTGGGTCC 135605
QY 430 rGluGlu-----PheLeuLysArgGluAlaArgAlaAsnGlnLeuAl 444
Db 135604 TGCTGGGAAGTCTTCCACTTTTCTCTCTTAAATCTTGCTCTCCACAAAGGTCTG 135545
QY 444 aGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspAspAr 464
Db 135544 TCCCAACCAGTCTAGCTCCCTGCCAGCAAGGTCTCTCAGTCTCTCTCTCTCTCCAGG 135485
QY 464 gSerGluGluLysTyrThrAla-----ValGlnArgAs 476
Db 135484 GAGCTACGCCCGTCTGTCTGAGCTCCAGGTCTCATCCCTGACCCCACTACAGACGTCC 135424
QY 476 nSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProG1 496
Db 135424 CAGGTCCGAGGCCCGGCTGG-----CCGG 135398
QY 496 yGlnArgAsnArgGluValIleSer----- 504
Db 135397 TGTCTCTTCACTGGAGGCGTGAACCTGTCCAGGTGGTGGCTGAGGCGAGGCCCTCTCGA 135338
QY 505 -----TrpGlySerGly-----ArgGlnAsnSe 512
Db 135337 CAGGCTCTCTCAGCACCCCTGGAGTGAGGGCTTGGGTCTTAGCCACACAAGGCTCTGA 135278
QY 512 rProArgMetGlyGlnProGlySerGlySerMetProSer----- 525
Db 135277 GCCCAGACAGGCTGAGNAGCTTCCGGGGCGGCTCCGCTCGTGGGGGGCTCTTTTAGGG 135218
QY 526 -----ArgSerThrSerHisThrSerAspPheAsnProAsn---SerGlySerAspG1 542
Db 135217 AGGGGACGGCCACTCCCGGAGGTGAGCACCTTCGGAGCAGAGTGTGGCGCTGCACCCA 135158
QY 542 nArgValValAsnGlyValProTrpProSerProCysProSerProSerProSerArgPr 562
Db 135157 AGCAGAGCAAGGGCGGTGAGCTTGGGTCTTGGGTGGCGGCCCGCCCATGCCACC 135098
QY 562 oProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrAr 582
Db 135097 TGGCAAGGTGGGATCTGGAAGTCTCTTCTACCC-----TGGGACCAGAA 135050
```

```
QY 582 gProProSerArgProProSerArg----- 590
:|||||
Db 135049 GCCCCCT-----CCTGCTTCAAGGATCTCGGGTGAGGAGGGGTGGCCAGAGTCCG 134996
QY 591 -----ProSerArgProProSerHisProSerHisAlaHisGlySe 603
:|||||
Db 134995 GAGGAGGGGTCTCTTCTCAACACGAGACTGGACCTGCACAGGTGGAGCTGGAAGGGC 134936
QY 603 rProAlaPro-ValSerThrMetProLysArgMetSerSerGluGlyProProArgMet- 622
:|||||
Db 134935 ACCACCTCCACCCAGGAGCATGGCAGGCGCACTCATTTGCATGCAGGCGGCACCGAGGCGAG 134876
QY 623 SerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIle 642
:|||||
Db 134875 AGC-----CGCCACCAACCCACCGCTGAAACGGATTGAGCAGAAAGCTCCGT 134828
QY 643 SerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAla----- 657
:|||||
Db 134827 TTCCTTCTGTCTGTCTCTCTCAAGGTCAAGCCCAAGCAGGAGCGGACAAAGGCGG 134768
QY 658 -----AlaThrProProValAlaAlaArgThrSerProSerGlyGlyThrTrpSer 673
:|||||
Db 134767 TGGGTCCCCCACTCAACCTCCCGCGCGCTGCGCTCTAGAGGAAGGACCTGGGA 134708
QY 674 SerValValSerGly-----ValProArgLeuSerProLysThrHisArgProArgSer 691
:|||||
Db 134707 AGGAGGGG-GAGGCGCAGGGCGCTCTCGAGCGCTGCGCCTGGCGCAGGCTGCCCGGCAAA 134649
QY 692 ProArgGlnAsnSerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAla 711
:|||||
Db 134648 CCC-----CTGCTGGCAGGGGGGCCACTGTGGGGGTCCAGAA 134610
QY 712 GlyIleIleProThrGluAlaValAlaMetProIleProAlaAlaSerProThrPro-Al 731
:|||||
Db 134609 GCCCTGTTCCTCTCTGGGGTGGAGCTGTCGGCTTTGGGAGCGAGCGCGCGGAG 134550
QY 731 aSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLysAspSerArgLeuG 751
:|||||
Db 134549 GTCACTCTAGAGTAAGCTCTCTGTGGAGCA-----AGGGTCAG 134511
QY 751 nAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSe 771
:|||||
Db 134510 AAATCAGGCTCACTCTCGG-----AACAGGTGTC 134481
QY 771 rProSerPheSerLysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLy 791
:|||||
Db 134480 GCCAAGTCAAGGGTTATCAGGGTCCCGCTCAGTCTCTGTCTGGAGCCAGTACAGGC 134421
QY 791 sGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSe 811
:|||||
Db 134420 T-----TGATGGGTCTGGCCCGCGCCCTGCTCT 134391
QY 811 rGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyLysSerArgAspLeuI 831
:|||||
Db 134390 GCACTGTGGGACTGGGGTCTTACCACGACTCAGAGGGA----- 134351
QY 831 eLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSerSerSerAsnCy 851
:|||||
Db 134350 -----CTGTCCAGCCCGTG-----CCAGCACCACCTG 134322
QY 851 sThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnTh 871
:|||||
Db 134321 CCCACATGGC----- 134312
QY 871 rGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCy 891
:|||||
Db 134311 -----CAGGCGCAGGCTTGGGATCAGCTCCCGCAGCCCTCAG 134271
QY 891 sLysGlnGluLysAspAspLysGluLysLysAsp-----AlaAlaGluGlnValArgLy 910
:|||||
Db 134270 CCTTCAAGGCGCGCAGCGAGGAGGAGCTTAAATAATGCGGGAGAAATTTTCATTAC 134211
QY 910 sSerThrLeuAsnProAsnAla-----LysGluPheAsnProArgSerPheSerGlnProLy 929
:|||||
```

```
Db 133262 C---CCTCCTCTCCTTCTCCTCTCCTCCTCCTTCCCTTCTCCTTCTCCTCTC----- 133214
Qy 1251 nSerGlyMetValProSerHisProThrAla-HisAlaPro-MetMetLeuMetThrThr 1270
Db 133213 -----TCCATCCTTCTCCTTCCATTCATCTCTTCTCCTCCACTCTG 133170
Qy 1271 GlnProProGlyGlyProGlnAlaLeuAlaGlnSerAlaLeuGlnProIleProVal 1290
Db 133169 TCCCTCCTCCTCCTCCTTCTCCTCTCCTT-----CTC 133137
Qy 1291 SerThrThrAlaHis-PhePro-----TyrMetThrHisProSerValGlnAl 1306
Db 133136 TCTCCTCATCCCACTTCAGTCTGGGGCTGAACACAGACCCGTTCTCCAGTGTCCAG 133077
Qy 1306 aHisHisGlnGln 1310
Db 133076 ACACCACCCAGCGT 133064

RESULT 11
US-11-121-086-6/c
; Sequence 6, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138-6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 172543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-6

Alignment Scores:
Pred. No.: 2,2e-05 Length: 172543
Score: 395.50 Matches: 364
Percent Similarity: 31.34% Conservative: 154
Best Local Similarity: 22.02% Mismatches: 586
Query Match: 5.68% Indels: 551
DB: 72 Gaps: 72

US-10-802-228-2 (1-1312) x US-11-121-086-6 (1-172543)
Qy 6 AlaAlaProArgSerProAlaValAlaThr--GluSerArgArgPheAlaAlaAlaArgT 25
Db 40817 GCGCGCGCGCTCACCTACCTGGGGTTCTCGGTTCCAGTAGACGGGTAGCGTCCGAGTT 40758
Qy 25 rpProGlyTrpArgSerLeuGlnArgProAlaArg----- 37
Db 40757 GCGCGCGCGCGGT-----CCTCGCGCGCGCGGAGGCGCGCGCGCGCGGTAAACG 40704
Qy 38 -----SerGlyArgGlyGlyGlyAlaAlaPro-GlyProTyrProSerAla 53
Db 40703 CAGGAGCAGCAGCGGAGCAGCGGCGGTGCGCGGCGCGCATGGCCCGGTCGCGCCGCC 40644
Qy 54 Ala-----ProProProGlyProGlyPro 62
Db 40643 GCGCTCGCTCTCCCGCGCGCGCTGCCAGCCTTCTCCGCTCTCTCCCGCGCGCGCGCG 40584
Qy 63 ProProSerArg----- 66
Db 40583 CCGCCGAG-CGGCGCGGCGGAGGCGGCGGAGGAGGCGCGCGCGCTTGGAGATCC 40525
Qy 67 GlnSerProProSerAlaSerAspCysPheGlySerAsnGlyAsnGlyGlyGlyAla 86
Db 40524 CGGCGCGCGCGCTCTCGCGCGCGCGACCTCGGCGCGCGAGAGTCAGCGCGCGCGG 40465
```

```
Qy 87 Phe-----ArgProGlySerArg-----ArgLeuLeuGlyLeuGlyGly 99
Db 40464 GCGGGGCGCGCGGGTTCGCGCGCGCGAGGACCGCGCGCGGACTGTCTCGAGCGCGG 40405
Qy 100 ProProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProAla 119
Db 40404 CCGCGCGCGCGCAC-----CCGCGCGGAGGGGCGCGCGG 40369
Qy 120 Ala-----ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArg--- 135
Db 40368 TCCGGCCTTTGTCGCGCGCGAGCGTCGGA---GGCGCGCGCGCGCCCTCCCGCCGTG 40312
Qy 136 -----SerGlyValSerLeuAlaArgProAlaProGlyCysPro 148
Db 40311 CCCCGCGCGCGGAGCTGGCGGGGTCTCCAGCCTGCGCGCGCGCCCTGCC----- 40258
Qy 149 ArgProAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysProGlnGln 168
Db 40257 ---CGGCG---CACCCCGTCTCAGG----- 40237
Qy 169 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPro 188
Db 40236 -----TCGTCCCGCGGGCCTGCCCGCAGCGACCCCGATTCTGCCTG 40195
Qy 189 ProProAlaAlaAlaAsnValArgLysProGly-----GlySerGlyLeuLeu--- 204
Db 40194 CCCCGCGCGCATCCGCGTGGCGCGCTGGGTCTCCGAGGGCGGGCGCACCCCTG 40135
Qy 205 -----AlaSerProAlaAlaAlaProSerProSerSer 216
Db 40134 CTTCCCGCGCGCCCGCGCTCCACCGCGCGCGCTGGGTCCCGCGCTCTCACCTGC 40075
Qy 217 SerValSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGly 236
Db 40074 CCGGTTCCAGCGCTATCCGAGCCTCACCAACCCCGAGCGCGCTCCCAATCGGGTC 40015
Qy 237 GlyArg-----ProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
Db 40014 GGCACACTCCAGTCTCCTCGCGGGCGCGGAGAGTTTCAGAGGGCGCTGGCCAAAG 39955
Qy 254 Ser-----ThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetVal 268
Db 39954 TCCAAAGCTGTCGCGGTGCGACCTCCTCGGGCGCTCTCGGAGTCGCTCTCGGACC 39895
Qy 269 HisIleLeuThr-----Ser 273
Db 39894 CACGCTGTCTCCTGGGGTGGAGTAAGTCCCTACACCTCAGCCCTAAGGTGCCGGAGC 39835
Qy 274 ValValGly-SerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPh 293
Db 39834 ATAGTAGGTGACCGATGTGGAGTGTGAGCAGCTGGGATGGGATCCAAAGAGCC----- 39780
Qy 293 eLysThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGl 313
Db 39779 -CAGAGTGAAGCCCG---GCCAGACGAAATCAGTCTCCACGACCAAGTCTCTGAGGCA 39724
Qy 313 uSerSerSerGlyProLysArgGluIleMetGlu-----SerIleLeuPhe--- 329
Db 39723 GGGGAGGGCGCTCCCGGGGCGACAGCTTGTGCAAGACCCGGGGTCTGATCTCTTCTAG 39664
Qy 329 ----- 329
Db 39663 GGAACGCGAGTGATTTCTCCTCTGTGTAGGTGACACAGGTGGGGAACCTGGGGCGCTTG 39604
Qy 329 ----- 329
Db 39603 CCTCTGCTTTGCTTTTTCATCTATAGACAGTGTCTGGCCATGTTGGCCACGCTG 39544
Qy 330 -----LysCysSerAspPh 334
Db 39543 GTCTTGAACTCCTGGGCTCAGGCGATCTTCCACCTGGGCTTCCAAAGTCTAGGATTA 39484
Qy 334 e-----ValValValGlnPh 339
```

Db 39483 CAGGCGTGAGCCACTGCACCTGGCGCTGCTGTTTGGCCCTTATTATCCACAGCC 39424  
QY 339 eLysAspMetAspSerSerTyAlaLys : : : : :  
Db 39423 TCCCTCACCAGACTCTGAGCTTGAGGCTAGGCTGTATCAAAAGCTGGCTGACATT 39364  
QY 349 : : : : : ArgAspAlaPheThrAsp : : : : : SerAl 356  
Db 39363 GCTGGCGATGCCAGATCGTGTGTACGTCTCTGAATTAGTGTATTCACAGGCCAGT 39304  
QY 356 alleSerAlaLysValAsnGlyGluHisLysGlu : : : : :  
Db 39303 GCAGGAGAGGAACTGCAGCGAAGCAGCAACAGCCCTTGGACCCGGAAGTCCCTGTGTG 39244  
QY 368 : : : : : LysAspLeuGluProTrpAspAlaGlyGluLeuThrAl 380  
Db 39243 AGGCATTACCAAACTCCTGGCAGGACCGTGAACCCAGAGATGGCAAGCTAACTGTG 39184  
QY 380 aLenGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMe 400  
Db 39183 CCCAGCAAGGTG : : : : : GGCTGGGACCGGACCCACC 39151  
QY 400 tPheArgTyAsnGluAsnTyr : : : : : GlyValValSerThrTyAspSerSe 417  
Db 39150 ACCGTCCACAGCTCTGAGACGTGGTGGCTGTCTGTGTCTCTGCTGCTCTCTC 39091  
QY 417 rLeuSerSer : : : : : TyrThrValProLeuGluArgAspAsnSe 430  
Db 39090 CCCCGATTCCCAGCGGCCACCGCGGCCACACCACTGACCTTGTTCAGGCTGGGTCC 39031  
QY 430 rGluGlu : : : : : PheLeuLysArgGluAlaArgAlaAsnGlnLeuAl 444  
Db 39030 TGCTGGGAAGTCTTCCACTTTTCTCTCTGTTAAATCTGTCTTCAGAAAGTCTG 38971  
QY 444 aGluGluLeuGluSerSerAlaGlnTyLysAlaArgValAlaLeuGluAsnAspAspAr 464  
Db 38970 TCCACCCAGTGTCTAGCGTCCCTCCAGGACAGGTTCTCAGTTCTCTCTCCCTCCAGG 38911  
QY 464 gSerGluGluGluLysTyThrAla : : : : : ValGlnArgAs 476  
Db 38910 GAGCTCAGCCCGTCTGTTGAGCTCCAGGTCTCATCCCTGACCCCACTACAGAGTCC 38851  
QY 476 nSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyRileProGln 496  
Db 38850 CAGGTGCAAGCCACAGGCTGG : : : : : CCGGG 38824  
QY 496 yGlnArgAsnArgGluValIleSer : : : : : 504  
Db 38823 TGTCTCTTCAGTGGAGGCGTGAACTGTCCAGTGTGTGGCTGAGGAGCCCTCTGGA 38764  
QY 505 : : : : : TrpGlySerGly : : : : : ArgGlnAsnSe 512  
Db 38763 CAGGCTCTTCAGACCCCTGGAGTGAGGCGCTTGGGTTCTAGCCACACAGGCTCTGTA 38704  
QY 512 rProArgMetGlyGlnProGlySerGlySerMetProSer : : : : : 525  
Db 38703 GCCCAGACAGCTGGAGAACTTCCCGGGCGCTCCGTCTGTGGGGGGCTCTTTAGGG 38644  
QY 526 : : : : : ArgSerThrSerHisThrSerAspPheAsnProAsn : : : : : SerGlySerAspG 542  
Db 38643 AGGGGACCGCCACTCCCGAGGTGAGCACTTTCGGAGCAGAGTGTGGCGCTGCACCCA 38584  
QY 542 nArgValValAsnGlyValProTrpProSerProCysProSerProSerSerArgPr 562  
Db 38583 AGCACAGGACAAGGGCCGTGAGCCTGGGTCTTGGCGTCCCGGCCCGCCAGATGCCACC 38524  
QY 562 oProSerArgTyGlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrAr 582  
Db 38523 TGCCAAAGGTGGGAGTCTCGAAGCTTCTCTTACCC : : : : : TGGGACACAGAA 38476  
QY 582 gProProSerArgProProSerArg : : : : : 590

Db 38475 GCCCCCT : : : : : CCTGCTTCAAGGATCTCGTGGTGAGGAGGCTGGCGCAGGCTCG 38422  
QY 591 : : : : : ProSerArgProProSerHisProSerAlaHisGlySe 603  
Db 38421 CAGGAGGGGTCTCTTCTCAAAACAGGACTGGACCTGCACAGGTGGAGCTGGAAGGCG 38362  
QY 603 rProAlaProValSerThrMetProLysArgMetSerSerGluGlyProProArgMet 622  
Db 38361 ACCACTCACCCAGGATGGCAGGCCACTCATTTGATGCAGCGGGCCACCGAGCGAG 38302  
QY 623 SerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIle 642  
Db 38301 AGC : : : : : CGCACCCACCGCCCTCAACCGATTTCAGCAGAGAAGCTCCGTT 38254  
QY 643 SerSerGlyLeuGluPheValSerHisAsnProProSerGluAla : : : : : 657  
Db 38253 TTCCTTCTGTTCTCTCTCAAGGTCAAGCCCAAGCAGGACGCGGACCAAGGCGG 38194  
QY 658 : : : : : AlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSer 673  
Db 38193 TCGTCCCCCATTCACCTCCCGCGCTGGCGCTCTCTAGAGAAAGGACCTGGGGA 38134  
QY 674 SerValValSerGly : : : : : ValProArgLeuSerProLysThrHisArgProArgSer 691  
Db 38133 AGGGAGGG-GAGGGCAGGGCGTCTCCGAGCCCTGGCCCTGGCGAGGCTGCCCGCAAA 38075  
QY 692 ProArgGlnAsnSerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAla 711  
Db 38074 CCC : : : : : CTGCTGCGAGGGGGCCACTCTGCGGGTCCCAAGAA 38036  
QY 712 GlyIleIleProThrGluAlaValAlaMetProIleProAlaAlaSerProThrProAl 731  
Db 38035 GCCCTGTTCCCTCCCTGGGCTGGAGCTGTCGCCGTTTGGACGCGCGCGCGGAG 37976  
QY 731 aSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLysAspSerArgLeuG 751  
Db 37975 GTCACTCTAGATGAAGTCTCTGTGGAGCA : : : : : AGGCTCAG 37937  
QY 751 nAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSe 771  
Db 37936 AAATCAGGCTCACTCTGGG : : : : : AACAGGTGTCC 37907  
QY 771 rProSerPheSerLysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLy 791  
Db 37906 GCCAAAGGTCAAGGTTATCCAGGCTGCCCTCAGTCCCTGGTCTGGAGCCAGTACAGGC 37847  
QY 791 sGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSe 811  
Db 37846 T : : : : : TGATGGGGTCTGGCCCGCCCTCTCTCT 37817  
QY 811 rGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyGlyLysSerArgAspLeuI 831  
Db 37816 GCACCTCTGGAGCTGGGGTCTACCCAGACTCAGGAAGGCA : : : : : 37777  
QY 831 eLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSerSerSerAsnCy 851  
Db 37776 : : : : : CTTGTCCAGCCCGTG : : : : : CCCAGCACCCACTG 37748  
QY 851 sThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnTh 871  
Db 37747 CCCACATGCG : : : : : 37738  
QY 871 rGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCy 891  
Db 37737 : : : : : CAGCCACAGGCTTGGATCAGCGTCCCGAGCCCTCAG 37697  
QY 891 sLysGlnGluLysAspAspLysGluGlyLysAsp : : : : : AlaAlaGluGlnValArgLy 910  
Db 37696 CCTTCAAGGCGCGCAGCGCAGGAGAAAGCTAAAAATATCGGGGAGAATTTCTCATAC 37637  
QY 910 sSerThrLeuAsnProAsnAla : : : : : LysGluPheAsnProArgSerPheSerGlnProLy 929  
Db 37636 AGAATAAATTTGGACCAATCGCTGCCCTGCCGCGAGCCTTTCAGAAAGCCCAT 37577

```
QY 929 sProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetVal-- 948
Db 37576 CTCAGCCAGGCAAGGATCAGGCGGAGGGGTACAGGGGCGCTGGGGCCCGGCGAGGTCA 37517
QY 949 -GlyHisGlnGln-----ProThrProValTyThrGlnPro----- 960
Db 37516 AGGCTCCACTCAGGCGGAGGAGCCACCGCTGTCCATCGCCCCCAGCTGGGAAGCAGCG 37457
QY 961 -----ValCysPheAlaProAsnMetMetTyProValProValSerProGlyValG1 978
Db 37456 AGCTGGGTCTGGCTTAGCCCGAG-----GGGCGAGGGGCG-- 37420
QY 978 nProLeuTyProLeProMetThrProMetProValAsnGlnAlaLysThrTyArgAl 998
Db 37419 -----TTTCCCTGGGTGGGTGGCTCGCTCTCTGAAATAAAATGTATCCCTTTCTCC 37367
QY 998 aValProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMet-HisP 1018
Db 37366 AGTGGGTAAATG-----AAAATCCCGAC 37343
QY 1018 rAlaSerAlaAlaGlyProProLeAlaAlaThrProProAlaTySerThrGlnTyrv 1038
Db 37342 CCCCATCAGCATGAGCCAGATTCTCGAACCCCTGCAGCCAGTGTGGAAAGAGC 37283
QY 1038 aAlaTySerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrg 1058
Db 37282 AGGCACACAGCCCC-----ATTCCAGCTCACCC---TCCTCCCTCTCCCGCAGCTCCC 37232
QY 1058 InSerGlnHisProHisValTySer-----ProValleGlnGlyA 1072
Db 37231 AFCCTTCCCATTTCTCTCTTTCACACTCTCTCCCTCCCTCCCTCCCTCCCTCCCT 37172
QY 1072 snAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAlaT 1092
Db 37171 CCTTCACACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 37112
QY 1092 hrGlnTyR-GlyAlaHisGlnGlnThrHisAlaMetTyAlaCysProLysLeuProTy 1111
Db 37111 CCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 37055
QY 1112 AnLysGluThrSer-----ProSerPheTyThrPhe 1121
Db 37054 TCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36995
QY 1122 AlaIleSerThrGlySerLeuAlaGlnGlnTyAlaHisProAsnAlaThrLeuHisPro 1141
Db 36994 CACCTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36938
QY 1142 HisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGly 1161
Db 36937 TTCACACCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36899
QY 1162 GlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAlaLeu 1181
Db 36898 ---TCCACACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36842
QY 1182 -----HisLeuAlaSerProGlnGlnGlnSerAla----- 1191
Db 36841 CTCCTCCCTTCCACACTCTCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36782
QY 1192 IleTyHisAlaGlyLeuAlaProThrProPro-SerMetThrProAlaSerAsnThrG1 1211
Db 36781 CTTTATACACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36722
QY 1211 nSerProGlnAnSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHi 1231
Db 36721 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36689
QY 1231 sValGlnProAlaTyThrAsnProProHisMetAlaHisValProGlnAlaHisValG1 1251
Db 36688 C---CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36640
```

```
QY 1251 nSerGlyMetValProSerHisProThrAla-HisAlaPro-MetMetLeuMetThrThr 1270
Db 36639 -----TCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36596
QY 1271 GlnProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProIleProVal 1290
Db 36595 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36563
QY 1291 SerThrThrAlaHis-PhePro-----TyMetThrHisProSerValGlnAl 1306
Db 36562 TCTCCCTCACCCACCTTCAGGTCTGGGCGCTGAACACAGACCCGCTTCTCCAGTGTCCAG 36503
QY 1306 aHisHisGlnGln 1310
Db 36502 ACACACACGAGGT 36490
RESULT 12
US-10-821-234-218/c
; Sequence 218, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 218
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-218
Alignment Scores:
Pred. No.: 1,56e-06 Length: 7006
Score: 394.50 Matches: 379
Percent Similarity: 31.31% Conservative: 143
Best Local Similarity: 22.74% Mismatches: 599
Query Match: 5.67% Indels: 557
DB: 6 Gaps: 74
US-10-802-228-2 (1-1312) x US-10-821-234-218 (1-7006)
QY 1 MetArgSerAlaAlaAlaPro-----ArgSerProAlaValAlaThrGluSer 17
Db 6977 TTGGCAGCAGCAGCAGCAGCAACACACAGCAGCAGCAGCGCGCGCGCGCGGAAACCCC 6918
QY 18 ArgArgPheAlaAlaAlaArgTrpProGlyTyTrpArgSerLeuGlnArgProAlaArgArg 37
Db 6917 GAGCGCCCGCGCGCAGCCCGCGCTTCCCGGA-----GCGCGAGCG 6879
QY 38 SerGlyArgGlyGlyGlyValAlaAlaPro-----GlyProTyPro-SerAlaAlaPr 55
Db 6878 GCGCGCAGCAGCAGCGGTCGCGCGCGCTTAGCTCGGCCCGCGCGCTCGGGGAC 6819
QY 55 oProProProGlyPro-----GlyProProSerArgGlnSerSerProProSe 72
Db 6818 CCCGACTCCCGCGCCAGCGAGCGCTCCCGCGCGCGCGCGCGAGAGCCGAGGAGGCGAGC 6759
QY 72 rAlaSerAspCysPheGlySerAsnGlyAsnGlyAlaPheArgProGlySerAr 92
Db 6758 GCGCGCAGAG-----CCCGGGGAGGGGGCGCGCA----- 6728
QY 92 gArgLeuLeuGlyLeuGlyGlyProArgProPheValVal----- 106
Db 6727 -----CCCGCGCGCGCGCGCATCTCTCAGGAGCCCGCAGAGCGCG 6690
```

QY 107 -----ValLeuLeuProLeuAlaSerProGlyAlaPro----- 117  
Db 6689 GAGGGCGCGGCGCGCGAGCGGTGCTGGCCCGCGGGCTCCCGGACTTCCCGCACCG 6630  
QY 118 -----ProAlaAla-----ProThrAr 123  
Db 6629 CTTGGGCGCGAGGAGCGCGTGTATCGGGCGGGCGCGCGCGCAAGGGTGGAGGAGCGG 6570  
QY 123 gAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeuAlaArgPr 143  
Db 6569 CCCCCCG 6510  
QY 143 oAlaProGlyCysProArgProAlaCysGluProValTyrglyProLeuThrMetSerLe 163  
Db 6509 GCGGCTGAAGCG 6451  
QY 163 uLySerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 183  
Db 6450 GCGGAGCG 6391  
QY 183 n-----GlnGlnGlnGlnProProProAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 198  
Db 6390 AGTGGCGGTACTGCGCAAGAGAGCATGGCCACAGCGCTTCTTCGTGCTGCGCGGAC 6331  
QY 198 roGlyGlySerGly-----LeuLeuAlaSerProAlaAlaAlaProSerPro----- 213  
Db 6330 CCGGCG 6271  
QY 214 -----SerSerSerValSerSerSerAlaThrAlaProSerSerValV 230  
Db 6270 TCAGTACTACGAGAGCGAGAAAGTGGCGGAGCAAGGCGCGCGCGCGCGCGCGCGCGG 6211  
QY 230 ala----- 231  
Db 6210 TCGTCTCGACTGCTGCTGAAATCAATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6151  
QY 231 ----- 231  
Db 6150 TCGCCTCTACACAGGACGAGTACTTCGCGTGGCGCGCGGAGAACGACGAGGACGAG 6091  
QY 232 -----AlaThrSerGlyGlyG 237  
Db 6090 AGGGCTGTACCG 6031  
QY 237 lyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSer----- 254  
Db 6030 CGCCCCCG 5971  
QY 255 --ThrIleSerPheAspGlyIleTyraAlaAsnMetArgMetValHisIleLeuThrSerV 274  
Db 5970 GCTCTGCG 5911  
QY 274 alValGlySerLysCysGluValGlnValLysAsnGlyIleTyrglyValPheL 294  
Db 5910 CCTACCGTGAAGTGTGCGAGGAGCACTGAGCGCGTGTGGTCCCGCGCGCGCGCGCGCG 5857  
QY 294 yThrTyrsrProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluS 314  
Db 5856 AGAACCTGACGGGGGTGACGCTGTGTGCGCTGTGTGCG----- 5819  
QY 314 erSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp- 333  
Db 5818 -----CGCACCATCGGCTTCGTGAAGCTCAACTGCGGAGCAGC 5782  
QY 334 --PheValValValGlnPheLysAspMetAspSerSerTyraAlaLysArgAspAlaPheT 353  
Db 5781 CGTCGGTGACCTGCAGCTCATATGAAATC-----CGCGCGTGGCGGCGCACT 5737  
QY 353 hrAspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProT 373  
Db 5736 CGGACAGCTTCTTCTATCGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5677  
QY 373 rp-----AspAlaGlyGluLeuThrAlaAsn-----GluGluL 384

Db 5676 GGATGCG 5617  
QY 384 euGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyraA 404  
Db 5616 TGAAGCGCGCTCAAG-----GAGCTCTTCAGTTCC 5587  
QY 404 snGluGluAsnTyrglyValValSerThrThrAspSerSerLeuSerSerTyThrValp 424  
Db 5586 GCGCGCGCG-----AGTAAGAGCAATCGTCGGGCTCGTCGCGCCAGCAC 5542  
QY 424 roLeuGluArgAspAsn-----SerGluG 432  
Db 5541 CCATCAGCTCCCG 5482  
QY 432 luPheLeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluLulleGluSerSer----- 450  
Db 5481 CCGGCGCTGTGGCGCGCTCGCGCACCGACAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 5422  
QY 451 --AlaGlnTyrlsAlaArgValAlaLeuGluAsnAspAspArgSerGluGluLyst 470  
Db 5421 GCAGCTCGTGGCG 5362  
QY 470 yThrAlaValGlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluA 490  
Db 5361 CCG 5310  
QY 490 snLysTyrlleProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgG 510  
Db 5309 -----GCCGCTGCG 5264  
QY 510 lnAsnSerProArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerH 530  
Db 5263 TCGCGCGCGCGCGCG-----GAGCAAGTGGCGCTCTCGCGCGCGCGCGCGCGCGCGCG 5207  
QY 530 isThrSerAspPheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValp 550  
Db 5206 AG-----C 5204  
QY 550 roTrpProSerProCysProSerProSerSerArgProProSerArgTyrglnSerGlyp 570  
Db 5203 CGCT-CCATGTCTATCGCGTGGCGACTCGCGCGCGCGCGCG-----CCACGAGCGCC 5154  
QY 570 roAsnSerLeuProProArgAlaAlaThrProThrArgProProSerArgProProSerA 590  
Db 5153 CGGCTCCCTGTCTCCAGCAGCGCGCGCGCG-----GCTCGGGCTCTTACCG-CCGCGCGCG 5098  
QY 590 rgProSerArgProProSerHisProSerAlaHisGlySerProAlaProValSerThrm 610  
Db 5097 GCCGCGACCGCGCTCTGCGCGATCCGCTGCACCGCGCG-----CCGCGC----- 5053  
QY 610 etProLysArgMetSerSerGluGlyProProArgMetSerProLysAlaGlnArgHisP 630  
Db 5052 -----AGCGCGCTCCAGCGCGCGCGCGCGCTCCG-----C 5025  
QY 630 roArgAsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValS 650  
Db 5024 CTCGG----- 5020  
QY 650 erHisAsnProProSerGluAlaAlaThrProProValAlaArgThrSerProSerGlyG 670  
Db 5019 -----GCTCCCG 4965  
QY 670 lyThrTrpSer-----SerValValSerGlyValProArgLeuSerProLysThrHisArgp 689  
Db 4964 CGACCTCGCGCGCTCTTCGAGCGCACCGAGCAACACCGCGCGAGTCCATCCGCGGAGCGCC 4905  
QY 689 roArg-----SerProArgGlnAsnSerIleGlyAsnThrPro-----SerGlyp 704  
Db 4904 CCGGCGCGGAGACG 4845  
QY 704 ro-----ValLeuAlaSerProGlnAlaGlylle-----lleProT 716

Db 4844 CCTGAGCCACTGTGGCGCTCTCTACCGCGGGTCTCTGGGGGACGCGCGCCAGGACCTGGA 4785  
Qy 716 hrGluAlaValalaMetProfileProAlaIaSerProThrProAlaSerProAlaSerA 736  
Db 4784 CCGAGGGCTGGCGAGA-----GGACCTACTCTCTCCAGCAGCCAGCCGCGCAGCG 4734  
Qy 736 snArgAlaValThrProSerSerGluAlaLysAspSer---ArgLeuGlnAspGlnArg 755  
Db 4733 GCCGTGCCCC-CAGCCCCCTCTCTGCTCGCTGGATGAATACACCCCTGATGCGGGCCACT 4675  
Qy 755 lnAsnSerProAlaGlyAsnLysGluAsnLysProAsnGluThrSerProSerPheS 775  
Db 4674 TCTCGGGCAGCGCGCGCTCTG-CGCGTCTGCCCCCGCTCTCTCCCAAGT----- 4620  
Qy 775 erLysAlaGluAsnLysGlyLeuSerProValValSerGluHisArgLysGlnIleAspA 795  
Db 4619 -----GGCCTACCACCCCTACCAGAGGACTACGGAGACATCGAGATC 4577  
Qy 795 spLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetA 815  
Db 4576 GGCTCCCAAG-----GAGCTCCAGCAGCAACCTCGGGGCGACAGCGCTTACATG 4526  
Qy 815 spGlnLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysI 835  
Db 4525 CCCATGACCGCGCGCGCTTCGCGGCGAGTGGGAGCGGAGCTGCAGGAGCGACAC 4466  
Qy 835 leGluProSerAlaLysAspSerPheIleGluAsnSerSerAsnCysThrSerGlyS 855  
Db 4465 TACATGCCCATGAGCCCCCGCAGCGTGTCCGCC-CCAAGCAGATCTTGCAGCCAGGGC 4407  
Qy 855 erSerLysProAsnSer-----ProSerIleSerProSerIleLeuSerAenThrG 872  
Db 4406 CGCGCCCGCGCC-GCCCGCGCGTCTCTGCGGGGCTGCG----- 4364  
Qy 872 luHisLysArgGlyProGluValThrSerGln----- 882  
Db 4363 -----GGGCCAGACCCACTCTTGGCGGGGCGAGGACATTCGCCGGCGAGCGGGG 4315  
Qy 883 --GlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAsp----- 897  
Db 4314 CGGGCTACAAGCGCAGCTCGCGCGGAGAGTCCCCCGGAGGACAGTGGGTACATGCCA 4255  
Qy 898 -----LysGluGluLysLysAspAla----- 904  
Db 4254 TGTGTGCGGTTCACAGGTCTCATGAGCATGTCAGATGGCAAGTGTGTCGCCAACGGG 4195  
Qy 905 -----AlaGluGlnValArgLysSerThr----- 912  
Db 4194 ACTACCTCAACGTGTCCCCAGCGAGCGGTTCACACCGGGCACCCTCCCGCTTCTTCT 4135  
Qy 913 -----LeuAsnProAsnAlaLysGluPheAsn----- 921  
Db 4134 CGCAGACCTGCACCCCGCGGGAGCGCTCAGGGGGTTCCTCCGGCTGCTGTCTACAGCT 4075  
Qy 922 -----ProArgSerPheSerGlnPro-----LysProSerThrT 933  
Db 4074 CTTTGGCCCCGTCTTACAAGGCCCTCTACACCTGTGTGGGGGCGACAGCAGCAGTACGTG 4015  
Qy 933 hrProThrSerPro-----ArgProGlnAlaGlnP 943  
Db 4014 TCATGAGCTTCCCCGTGGGGCGCATCTCTGGAGGAGGCGTCTGGAGCCTCAGGCGCAGC 3955  
Qy 943 roSerProSerMetValGlyHis-----GlnGlnProThrProV 956  
Db 3954 CAGGGCCCGCAGCGCGCGCAGCGCTTCCTGGGGCGCGCCCCCAGCAGCCCCCTCACCTG 3895  
Qy 956 alThrThrGlnProVal---CysPheAlaProAsnMetMetTyrProValProValSerPr 975  
Db 3894 TAGTGCCCTTCGCGCGCGGTAGCGCGCGCGGTTCAGTTTCTTGGGGCCAGCGCGGCC 3835  
Qy 975 oglyValGlnProLeuTyrProIlePro---MetThrProMetProValAsnGlnAlaLy 994  
Db 3834 GGGCGGTGAGGCCACGCGCTGTCTCTCTGGAGGGGCTGCCACGCTGCCCGCATGCGACG 3775

Qy 994 sThrTyArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAl 1014  
Db 3774 AGTACCCACTGCCACCGGAGCCCAAGAGCCCCGGGAGTACATCAACATCGACTT----- 3720  
Qy 1014 aMetMetHisProAlaSerAlaAlaGlyProProIleAlaAla-ThrProProAlaTyrS 1034  
Db 3719 -----TGGCGAGCCCGGGGCCCGCTGTGCGCCCGCGGCTCTCCCTGTGTGG 3673  
Qy 1034 erThrGlnTyrValAlaTyrSer-ProGlnGlnPhePro-AsnGlnPro---LeuVal 1051  
Db 3672 CGTCG-----GCGGCTCTCTCTCTCTCTCTGTCGCCAGCAGCCCGGCTCTGTCG 3619  
Qy 1052 -----GlnHisValProHisTyrGlnSerGln----- 1060  
Db 3618 TGGGCTCAGGCACCCCGGGCACACGAGCAGCAGCGGCGAGCGGTCTCCGCTCTCCGACT 3559  
Qy 1061 ----HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaPro P 1079  
Db 3558 ACATGAACCTCGACTTCAGCTCCCCCAAGTCTCTTAAGCCGGGCGCCCCGAGCGGCCACC 3499  
Qy 1079 roThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluG 1099  
Db 3498 CCGTGGGCTCTTGAGCGCTCTCTGTCTCCCGGAGGCTCTCTCC----- 3455  
Qy 1099 lnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerP 1119  
Db 3454 -----CCGTATCCGCGCTTGCCTCCCGCGCTCTCCCGCTCTCCCGCTCG- 3413  
Qy 1119 heTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrL 1139  
Db 3412 -----TCGTCTCT-CCAGCGCGCGCACCGCGCGCGCCCCCGGGGAGCTG 3368  
Qy 1139 euHisPro-----HisThrProHisPro----- 1146  
Db 3367 TACGCGCTGCCCGCGCTCGGCGCTTGCACCCCGCAGGCCCGGGGCGCTCTCATCG 3308  
Qy 1147 ----- 1147  
Db 3307 TTGTCTCGGACACCGGGGACAATGGTGACTACACCGAGATGGCTTTTGGTGTGCGGCC 3248  
Qy 1147 lnProSerAlaThr-----ProThrGlyGln----- 1155  
Db 3247 ACCCGCGCGCAACCTATTCGCGGCCCGCCCCGAGCAGAGCTGCCCGCTGCCAGCCG 3188  
Qy 1156 -----GlnGlnSerGlnHisGly----- 1162  
Db 3187 AGTCTGGGCGTGAAGAGGCTGAGCTCATGGAGCAGGTGTCTGGAGTCTGAGGCTTCTCTG 3128  
Qy 1162 lySerHisProAlaProSerProValGlnHisHisGln---HisGlnAlaAlaGlnAlaL 1181  
Db 3127 CAGGCCAGCCAGCCCGGACCCACCGCGCGCCCAAGGTTCATCCGCGCAGACCGCGCAG 3068  
Qy 1181 euHisLeuAlaSerProGln---GlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProT 200  
Db 3067 GGGGGCGCGCCCGCCAGTTCGAGACCTTCTCTCTCCACAGCAGCGTCA-CCCCCGT 3009  
Qy 1200 hrPro---ProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProA 1219  
Db 3008 GTCCCCGTCTCTCCCGCACACCCCAAGCGGCCAACATCTGGGCTCTCCGTGGGAAATGTCTC 2949  
Qy 1219 laAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPr 1239  
Db 2948 TCTCAGGAAAGCAGCAGGCGCGGTGGTGTGGCCCTTGAGGGGGCGCAGAGCGGCC 2889  
Qy 1239 roPro-----HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValP 1256  
Db 2888 CACCTCCCCACAGCATGTTGACGGCGCGCCCTTGTGCAACCGAGGCGCGCGCTGGAC 2829  
Qy 1256 ro----- 1262  
Db 2828 CCGGGTTCAGCCCGGGGCTTGGTGTCTTCTGGAGCGGTGGATCGCCCATCGCGCAG 2769

```
QY 1262 isAlaProMet-----MetLeuMetThrThrGlnPro----- 1272
D 1262 isAlaProMet-----MetLeuMetThrThrGlnPro----- 1272
Db 2768 AGAGACCTCTGCGGCTTCCAGATGCTCTCAACTACATCGCATCGAGGAGGA 2709
QY 1272 ----- 1272
Db 2708 GCCCGGCTGCCACCCAGCGCGCGCGCGCGCTTCTCAGCCGGGAGACAA 2649
QY 1273 -----ProGly---GlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGln----- 1286
Db 2648 GAGCTCTGCGGCGCGGACCGAAGCCTCGGGGTCTCATCAGCGCTGTGGGGTTCGGCAG 2589
QY 1287 -----ProLleProValSerThrT 1293
Db 2588 CACCGCGCGGGTGTGGGGGGCGGGTCCCGGTGCGCCTGCGCCCTGCCCGCTGCCAACACCTACGC 2529
QY 1293 hrAla 1294
Db 2528 CAGCA 2524

RESULT 13
US-11-087-100-1
; Sequence 1, Application US/11087100
; Publication No. US2005026640A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Platt, James
; APPLICANT: Kuner, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
; TITLE OF INVENTION: System and Uses Thereof
; FILE REFERENCE: 2997-29
; CURRENT APPLICATION NUMBER: US/11/087,100
; CURRENT FILING DATE: 2005-03-21
; PRIOR FILING DATE: 09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/284,066
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 60/298,796
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/323,269
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 8730
; TYPE: DNA
; ORGANISM: Schizochytrium sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8730)
US-11-087-100-1

Alignment Scores:
Pred. No.: 4.02e-06 Length: 8730
Score: 386.00 Matches: 295
Percent Similarity: 33.26% Conservative: 177
Best Local Similarity: 20.79% Mismatches: 507
Query Match: 5.55% Indels: 444
DB: 7 Gaps: 53

US-10-802-228-2 (1-1312) x US-11-087-100-1 (1-8730)
QY 2 ArgSerAlaAlaAlaProArgSerProAlaValAlaAlaThrGluSerArgArgPheAla 21
Db 1223 CGCTCAAGCACAGACTCTCCCGGGCCACATCAACGTGCGACACCCCAACCTCTACG 1282
QY 22 AlaAlaArg-----TrpProGlyTrp 28
Db 1283 ACAACACGCCCATCAACGAGTCTCGCTCTACATTAACACATGAAACCGCCCTGGTTC 1342
QY 29 ArgSerLeuGlnArgProAlaArgArgSerGlyArgGlyGlyGlyAlaAlaProGly 48
Db 29 ArgSerLeuGlnArgProAlaArgArgSerGlyArgGlyGlyGlyAlaAlaProGly 48
```





```
Db 3995 CGCCCTCTGCTCTGCAAGAGCTCTTTGAGAGGCCGAGACCGCTGCTCATGGAGTCC 4054
Qy 1008 AspGlnHisGlnSerAlaMetMet-----HisProAlaSerAlaGlyPro 1024
Db 4055 TCGCCGCCAAGACTGGCTACGAGCTGACATGATCGAGTCCGACATGGAGCTCGAGACCG 4114
Qy 1025 ProIleAlaAlaThrProProAlaThrSerThrGlnTyrValAlaTyrSerProGlnGln 1044
Db 4115 AGCTCGGCATGTGATCCATCAAGCGTGTG-----AGATTCTCTCGAGGTCC 4162
Qy 1045 PhePro-----AsnGlnProLeuValGlnHisVal----- 1054
Db 4163 AGGCCATGTCTCAAGCTCGAGGCCAAGAGCTGCGACGCTCTCAGCCGCCACCGCAGCTGTG 4222
Qy 1055 -----ProHisTyrGlnSerGlnHisProHisValTyrSer 1066
Db 4223 GCGAGCTCGTCGATGCCATGAAGCCGAGATCGCTGGTGGCTCTGCCCGGCGCTGCCG 4282
Qy 1067 ProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeu 1086
Db 4283 CGCTGCTC-----CTGCTCCGCTGCTCGCGCCCTGCGCTGCGCCCTGCGC 4333
Qy 1087 valSerSerSerAlaThrGlnTyrGlyAlaHisGlnThrHisAlaMetTyrAlaCys 1106
Db 4334 CTGCTCTCTCGAGC----- 4363
Qy 1107 ProLysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGly 1126
Db 4364 CGAGACTGTGTCATGGAGTCTCTGCCGCCAAGACTGGCTAGGACTG-----ACATGA 4420
Qy 1127 SerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisPro----- 1141
Db 4421 TCAGTCCGACATCGAGCTCGAGACCGAGCTCGGCATTGACTCCATCAAGCGGTGCGAGA 4480
Qy 1142 HisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGly 1161
Db 4481 TTCTCTCCGAGGTCGAGCCATCTCAACG----- 4510
Qy 1162 GlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAlaLeu 1181
Db 4511 ---TCGAGGCCAAGGAGCTCGAGCTC-----TCAGCCGACCGGCACTG 4552
Qy 1182 HisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrPro 1201
Db 4553 TTGCGCGAGTGTGTC-GATGCCATGAAGCCGAGATCGCTGGTCTGCTGCCCGCGGCT 4611
Qy 1202 ProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGln 1221
Db 4612 GCGCGCGCTGCTGCTGCTCGGCTGTGTCGCGCCCTGCGCTGCGCCCTGCG----- 4665
Qy 1222 GlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsn 1238
Db 4666 -----CCTGCCGCCCTGCGCTGCTGCTCGAGC 4695

RESULT 14
US-11-087-084-1
; Sequence 1, Application US/11087084
; Publication No. US20050273883A1
; GENERAL INFORMATION:
; APPLICANT: Metz, James
; APPLICANT: Barclay, William
; APPLICANT: Platt, James
; APPLICANT: Kuner, Jerry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide Synthase
; FILE OF INVENTION: System and Uses Thereof
; FILE REFERENCE: 2997-29
; CURRENT FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US/11/087,084
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: 09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/284,066
; PRIOR FILING DATE: 2001-04-16
```

```
; PRIOR APPLICATION NUMBER: 60/298,796
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/323,269
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 8730
; TYPE: DNA
; ORGANISM: Schizochytrium sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8730)
US-11-087-084-1

Alignment Scores:
Pred. No.: 4,02e-06 Length: 8730
Score: 386.00 Matches: 295
Percent Similarity: 33.26% Conservative: 177
Best Local Similarity: 20.79% Mismatches: 507
Query Match: 5.55% Indels: 444
DB: 7 Gaps: 53
```

US-10-802-228-2 (1-1312) x US-11-087-084-1 (1-8730)

```
Qy 2 ArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPheAla 21
Db 1223 CGTCAAGCAAGACTCTCCGGGACCATCAACGTCGACCAACCCACCACTTACG 1282
Qy 22 AlaAlaArg-----TrpProGlyTyr 28
Db 1283 ACAACAGCCCATCAACGAGTCTCGCTCTACATTAAACACCATGAACCGCCCTGGTTCC 1342
Qy 29 ArgSerLeuGlnArgProAlaArgSerGlyArgGlyGlyAlaAlaProGly 48
Db 1343 CGCCCTCTGGTGGCCCGCGCGCATTTTCGAGCTTTGGCTTTGGTGGCGCAACT 1402
Qy 49 ProTyrProSerAlaAlaProPro----- 56
Db 1403 ACCACCGCTCTCGAGGAGCGCCGACCGCAGCACACCGCGTACCGCTCAACAGC 1462
Qy 57 -----ProGlyProGlyPro----- 62
Db 1463 GCCCGAGCCGCTGCTCATGATGGCCGCCACCGCCCGGCGCTCCAGTCTGCGAGG 1522
Qy 63 ProProSerArgGlnSerSerPro----- 71
Db 1523 CCAGCTCAAGGAGTTCGAGGCGCCCATCAAGGAGAACGAGACCGTCAAGAACACCGCCT 1582
Qy 72 ---SerAlaSerAspCysPheGlySerAsnGlyAsnGlyGlyAlaPheArgProGly 90
Db 1583 ACATCAAGTGGGTCAAGTTTCGGCGAGCAGTTCAAAATTTCCTGGCTCCATCCCGGCCCAA 1642
Qy 91 SerArg-----ArgLeuLeuGlyLeuGlyGlyProArgProPhe 104
Db 1643 ACGCGCGCTCGGCTTCTCTCGTCAAGGATGCTGAGGATGCTGCTCCACCTCGTGCCA 1702
Qy 105 ValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAlaProThrArgAla 124
Db 1703 TCTGCGCCCAATTTCGCAAGGATGTCAACCAAGAGGCGCTGGCGCTCCCGCGAGGCG 1762
Qy 125 SerProLeuGlyAlaArgAlaSerProProArg----- 135
Db 1763 TCAGTTTCGCGCGCAAGGCGCATCGCCACCAACCGCGCTGTGCGCGCTCTTCTCCGCGC 1822
Qy 136 -----SerGlyValSerLeuAla 141
Db 1823 AGGCGCGCAGTACACGCACATGTTTTCGAGGAGTGGCCATGAACCTGGCCCGCTCGCC 1882
Qy 142 ArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThrMet 161
Db 1883 AGAGCATTCGCCCATGGAGCGCCCGCCCGCTGCTGCAAGGTCGCTGAGGACGACAGCTTG 1942
```



Db 3501 CCAGGCCATGCTCAATGTCGAGCCCAAGGATGTC-GATCCCTCTCAGCCCACTCGCACTG 3559  
Qy 841 spSerPheIleGluAenSerSerAsnCyeThrSerGlySerSerIysProAsnSerP 861  
Db 3560 TTGGTGAGTTGTCAACGCCCATGAAGCCGAGATCGCTGCAGCTCTGCCCGCGCGCTG 3619  
Qy 861 roSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrS 881  
Db 3620 CTGCGCTCTCGGCTCGCGGCAAGCTGCGCCCTGCGCCGCTGCGCTGCTGTCTCGA 3679  
Qy 881 exGlnGlyValGlnThrSerSerProAlaCysLysGln----- 893  
Db 3680 ACAGCTTCTCGAAGAGCCGAGACCGCTCGTATGGAGCTCTCGCGCCCAAGACTGGCT 3739  
Qy 894 -----GluLysAspLysGluGluLysLysAspAlaGluGlnV 908  
Db 3740 ACAGACTGACATGATCGAGTCGACATCGAGCTCGAGCTCGAGCTCGGCATTGACTCCA 3799  
Qy 908 alArgLys-SerThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGln 927  
Db 3800 TCAAGCGTGTGAGATCTCTCCG-----AGGTTCAAGGCCATGCTCAACGTCGAGG 3850  
Qy 928 ProLysProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMet 947  
Db 3851 CCAAGGACGTGCGACGCTCTCAGCCGCACTCGCACTGTGGTGAGGTGCTCA----- 3901  
Qy 948 ValGlyHisGlnGlnProThrProValTyrThrGlnProValCysPheAlaProAsnMet 967  
Db 3902 -----ACGCCATGAAGCTGAGATCGCTGTGTGCTCTGCCCGCGCG 3943  
Qy 968 MetTyrProValProValSerProGlyValGlnProLeuTyrProIleProMetThrPro 987  
Db 3944 CTGCCCGCGCTG-----CCCCAGGTCTGCTGTGCGCCCTGCGCTGCGCGCG 3994  
Qy 988 MetProValAsnGlnAlaLysThrTyrArgAlaValProAsnMetProGlnGlnArgGln 1007  
Db 3995 CGCCCTCTGTCTCGAAGAGCTTCTTGAGAAGCCGAGACCGTGTGTGATGGAGTCC 4054  
Qy 1008 AspGlnHisGlnSerAlaMetMet-----HisProAlaSerAlaAlaGlyPro 1024  
Db 4055 TCGCCCCCAAGACTGCTACGACTGACATGATCGAGTCCGACATGGAGCTCGAGACCG 4114  
Qy 1025 ProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGln 1044  
Db 4115 AGCTCGCATTTGACTCCATCAAGCGTGTCTG-----AGATTCTCTCGAGGTCC 4162  
Qy 1045 PhePro-----AsnGlnProLeuValGlnHisVal----- 1054  
Db 4163 AGCCCATGTCAACGTGAGGCCCAAGGACGTGCGAGCTCTCAGCCGACCCGCACTGTTG 4222  
Qy 1055 -----ProHisTyrGlnSerGlnHisProHisValTyrSer 1066  
Db 4223 GCAGGTGCTCGATGCCATGAAGCCGAGATCGCTGTGTGCTGTGCCCGCGCTCGCG 4282  
Qy 1067 ProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeu 1086  
Db 4283 CCGCTGCTC-----CTGCTCCGCTGCTGCGCCCTGCGCTGCGCCCTGCGCG 4333  
Qy 1087 ValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCys 1106  
Db 4334 CTGCTGTCTCGAGCG-----AGCTTCTCGAAGG 4363  
Qy 1107 ProLysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGly 1126  
Db 4364 CCAGACTGTGTCATGGAGTCTCGCCGCAAGACTGCTGCTACGAGACTG---ACATGA 4420  
Qy 1127 SerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisPro----- 1141  
Db 4421 TCGAGTCCGACATGGAGCTCGAGACCGAGCTCGGCATTGACTCCATCAAGCGTGTGAGA 4480  
Qy 1142 HisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGly 1161  
Db 4481 TTCTCTCCGAGGTTCAGGCCATGCTCAACG----- 4510

Qy 1162 GlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAlaLeu 1181  
Db 4511 ---TCAGGCCCAAGACGCTCGACGCTC-----TCAGCCGACCCGCACTG 4552  
Qy 1182 HisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrPro 1201  
Db 4553 TTGGCGAGTGTCTC-GATGCCATGAAGCCGAGATCGCTGTGGTGTCTGCCCGCGCGCT 4611  
Qy 1202 ProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGln 1221  
Db 4612 GCCGCGCTGTCTCTGCTCGCGCCCTGCGCCCTGCGCCCTCGC----- 4665  
Qy 1222 GlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsn 1238  
Db 4666 -----CCTCCGCGCCCTGCGCTGTCTGTCTCGAGC 4695

## RESULT 15

US-11-087-085-1  
; Sequence 1, Application US/11087085  
; Publication NO. US20050273884A1  
; GENERAL INFORMATION:  
; APPLICANT: Metz, James  
; APPLICANT: Barclay, William  
; APPLICANT: Platt, James  
; APPLICANT: Kuner, Jerry  
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORPA of a PUFA Polyketide Synthase  
; FILE REFERENCE: 2997-29  
; CURRENT APPLICATION NUMBER: US/11/087,085  
; CURRENT FILING DATE: 2005-03-21  
; PRIOR APPLICATION NUMBER: 09/231,899  
; PRIOR FILING DATE: 1999-01-14  
; PRIOR APPLICATION NUMBER: 60/284,066  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 60/298,796  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/323,269  
; PRIOR FILING DATE: 2001-09-18  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 1  
; LENGTH: 8730  
; TYPE: DNA  
; ORGANISM: Schizochytrium sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(8730)  
US-11-087-085-1

Alignment Scores:  
Pred. No.: 4,02e-06 Length: 8730  
Score: 386.00 Matches: 295  
Percent Similarity: 33.26% Conservative: 177  
Best Local Similarity: 20.79% Mismatches: 507  
Query Match: 5.55% Indels: 444  
DB: Gaps: 53

US-10-802-228-2 (1-1312) x US-11-087-085-1 (1-8730)

Qy 2 ArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPheAla 21  
Db 1223 CGCTCAAGCACACAGACTCTCCCGGCGACCATCAACGTCGACACACCCCAACTTACG 1282  
Qy 22 AlaAlaArg-----TrpProGlyTyr 28  
Db 1283 ACACAGCCCATCAACGAGTCTCTACATTAAACACCATGAACCCCGCTGTGTTC 1342  
Qy 29 ArgSerLeuGlnArgProAlaArgArgSerGlyArgGlyGlyAlaAlaProGly 48  
Db 1343 CGCCCCCTGTGTGTGCGCCCGCGCGCATTTTCGAGCTTTGCTTGTGGTGGCCCAACT 1402  
Qy 49 ProTyrProSerAlaAlaProPro----- 56





**This Page Blank (uspto)**